

国际信息工程先进技术译丛

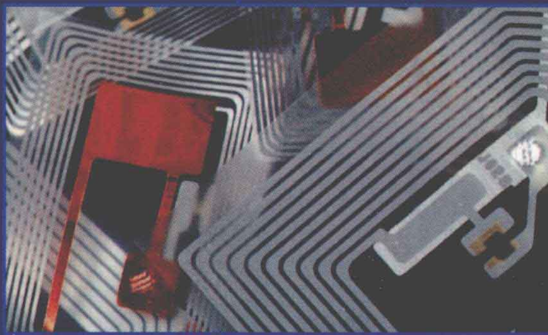
CRC Press
Taylor & Francis Group

RFID与传感器网络：架构、协议、安全与集成

RFID and Sensor Networks: Architectures, Protocols, Security and Integrations

（挪）Yan Zhang
（加）Laurence T. Yang　编著
（中）Jiming Chen
谢志军　等译







机械工业出版社
CHINA MACHINE PRESS

国际信息工程先进技术译丛

RFID与传感器网络：架构、协议、安全与集成

（挪） Yan Zhang
（加） Laurence T. Yang 编著
（中） Jiming Chen
谢志军 等译



机 械 工 业 出 版 社

传感器技术、微机电系统、微电子技术和无线通信等技术的进步，推动了无线传感器网络与RFID的产生和发展。

本书为读者提供了一个综合的技术指导，全书分为三大部分，第1部分介绍RFID的基本原理，例如，标签、阅读器、中间件、安全和服务。第2部分介绍WSN的基本原理，例如，路由、媒介访问控制、定位、聚类、移动性、安全和跨层最优化。第3部分探究将RFID与WSN集成的规则和应用。

本书可作为通信、计算机类专业的高年级本科生和研究生教材，也可供相关专业的科学研究工作者和工程技术人员参考。

图书在版编目(CIP)数据

RFID与传感器网络：架构、协议、安全与集成/(挪)张彦，(加)杨(Yang, LT.)，(中)陈积明编著；谢志军等译．—北京：机械工业出版社，2012.
(国际信息工程先进技术译丛)
RFID and Sensor Networks: Architectures, Protocols, Security and Integrations
ISBN 978-7-111-37399-5

Ⅰ.①R… Ⅱ.①张…②杨…③陈…④谢… Ⅲ.①无线电信号—射频—信号识别②无线电通信—传感器 Ⅳ.①TN911.23②TP212

中国版本图书馆 CIP 数据核字(2012)第077245号

机械工业出版社（北京市百万庄大街22号　邮政编码 100037）
策划编辑：林　桢　责任编辑：林　桢
版式设计：刘怡丹　责任校对：张晓蓉
封面设计：马精明　责任印制：乔　宇
北京瑞德印刷有限公司印刷（三河市胜利装订厂装订）
2012年7月第1版第1次印刷
169mm×239mm · 32印张 · 660千字
0001—2500册
标准书号：ISBN 978-7-111-37399-5
定价：138.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务	网络服务
社服务中心：(010)88361066	门户网：http：//www.cmpbook.com
销售一部：(010)68326294	教材网：http：//www.cmpedu.com
销售二部：(010)88379649	
读者购书热线：(010)88379203	**封面无防伪标均为盗版**

译 者 序

传感器技术、微机电系统、微电子技术和无线通信等技术的进步，推动了无线传感器网络（WSN）与射频识别（RFID）技术的产生和发展。RFID具有广阔的应用前景，可以应用于供应链管理、电子支付、环境监控、办公室访问控制、智能标签、目标探测与跟踪、港口管理以及食品工业监控等方面，同时也是实现物联网的基础。而无线传感器网络集数据的采集、传输、融合和分析于一体，是信息技术的一个新领域，它作为连接物理世界和网络虚拟世界的桥梁，拓展了人们获取信息的能力，能够将客观世界的物理信息同传输网络连接在一起，将为人们提供最直接、最有效和最真实的信息。无线传感器网络具有广阔的应用前景，能应用于环境监测、城市管理、生物医疗、卫生保健、抢险救灾、工农业控制、危险区域过程控制、防恐反恐和国防军事等领域。从20世纪90年代末开始，以美国加州大学伯克利分校、加州大学洛杉矶分校等为代表的各国研究机构纷纷开展了无线传感器网络的理论和实验研究。美国《技术评论》（Technology Review）杂志于2003年2月份评选出对人类未来生活产生深远影响的十大新兴技术，无线传感器网络位列第一。随着传感器网络的出现，人们获取数据的方式经历了从手工到自动化，从有线到无线，收集范围从日常生产生活的区域到许多前所未及的自然空间；这些改变使我们获得了更多的数据，而从这些数据中提取的有用信息，正逐渐地改变我们所处的世界。由于RFID与WSN具有技术互补、组合灵活的特性以及普适计算的特点，因此将这两种技术集成是未来技术的一个发展趋势。

本书为读者提供了一个综合的技术指导，全书分为三大部分，第1部分介绍RFID的基本原理，这部分让读者对RFID有一个了解，例如标签、阅读器、中间件、安全和服务。第2部分介绍WSN的基本原理，这部分让读者对WSN有一个了解，例如路由、媒介访问控制、定位、聚类、移动性、安全和跨层最优化。第3部分探究将RFID与WSN集成的规则和应用。

本书可作为通信类专业、计算机类专业和电子类专业的高年级本科生和研究生教材，也可供相关专业的科学研究工作者和工程技术人员参考。本书的翻译和出版得到了国家自然科学基金项目（60902097）、科技部公共服务平台创新基金(09C26243314159)、国际科技合作项目制（2009DFA12120）、浙江省重大科技专项重点工业项目（No. 2011C11042）、浙江省自然科学基金项目（Y1090571）和宁波大学胡岚博士基金的支持，在此谨致谢意。本书的翻译工作由宁波大学谢志军老师主持，吕玲红、沈怡飞、何伟、于凯和杨方清分别参与了本书的翻译工作，特此向支持和关心本书翻译工作的所有单位和个人表示衷心的感谢。特别感谢宁波大学

信息学院研究生吕玲红和宁波大学外语学院的沈怡飞为本书所做的翻译、排版和校译工作。还要感谢译者的家人，教育译者多年的师长、学长和同仁的帮助和支持。由于本书研究的内容属于信息科学研究的热点和前沿科学问题，许多专业术语的中文译文还没有统一的标准，加之译者水平有限，虽几经修改，书稿翻译过程中错误和不足在所难免，敬请读者指正。

谢志军

2012.6

原书前言

近来，射频识别（RFID）技术在工业界和学术界得到迅速发展。供应链管理、电子支付、RFID护照、环境监控、办公室访问控制、智能标签、目标探测与跟踪、港口管理、食品工业监控以及动物鉴定等方面都将应用到这一技术；RFID也是实现普适计算——物联网的基础。我们完全有理由相信：当RFID的原理得到普遍理解、元器件成本得到降低RFID的安全性得到保障的时候，更多的基于RFID的产品将会出现。

随着数字电子技术、嵌入式系统、信号处理以及无线通信等技术的发展，人们也开始关注无线传感器网络（WSN）。无线传感器网络是由大量体积微小的，带有感知、控制、数据处理以及通信和联网能力的传感器节点组成的。WSN具有如下一些特点：网络中节点部署密集、传感器节点通常不可靠、网络的拓扑频繁变化和节点的能量、计算和内存严重限制。另外，传感器网络通常在恶劣的无人监护的环境下进行数据传输，它的需求与设计明显地不同于传统的蜂窝网络、自组织网络或者网状网络。这些特性给WSN的设计带来了诸多挑战。

在实际的应用中，由于RFID与WSN具有技术互补、组合灵活的特性以及普适计算的特点，将这两种技术集成是未来技术的一个发展趋势。例如在智能家居、监控系统和个人医疗等应用中就已经将这两种技术进行了有机的集成。这两种互补的技术的有机整合能够呈指数级地提高应用监控能力。然而，与单独采用RFID或者无线传感器网络技术相比，集成RFID和WSN这两种技术面临更多的技术、操作、商业以及政策方面的挑战。

本书为读者提供了一个综合的技术指导，包括基本概念、基础技术、最新进展，同时也分享了RFID和WSN以及集成技术中的问题或争议；本书结合图表进行讲述，并且能够进行完整的交叉对照。另外，本书还提供了一些详细和特定的用于提高RFID、WSN以及集成系统效率的技术信息。

本书由三部分组成。

第1部分：RFID

第2部分：WSN

第3部分：RFID与WSN集成

第1部分介绍RFID的基本原理，这部分让读者对RFID有一个了解，例如标签、阅读器、中间件、安全和服务。第2部分介绍了WSN的基本原理，这部分让读者对WSN有一个了解，例如，路由、媒介访问控制、定位、聚类、移动性、安全和跨层最优化。第3部分探究将RFID与WSN集成的规则和应用。

本书有如下特色。

1）可以作为RFID、WSN以及其集成技术的全面的必不可少的参考工具。

2）包含了基本原理、系列概念以及未来发展方向。

3）介绍了架构、协议、标准、安全及其应用。

4）帮助专家、工程师、学生以及研究员更进一步地理解RFID和WSN。

5）提供了RFID与WSN集成方面独树一帜的内容。

本书可以作为无线通信和网络技术领域的学生、教育工作者、研究决策人员、科研工作者以及工程技术人员的参考书和工具书，尤其能够吸引那些正在开发RFID、WSN及其集成应用的学生、研究员、开发人员和顾问，并得到他们的青睐。

我们特别感谢所有那些为本书付出辛勤劳动与时间的撰稿者，感谢他们为本书做出的杰出贡献，所有人都表现出了极其敬业和高度的合作精神。特别感谢Taylor & Francis Group的Richard O'Hanley, Stephanie Morkert和Joette Lynch的支持，感谢他们自始至终的耐心和专业精神。我们很感激Sridharan Sathyanarayanamoorthy辛劳的排版工作。最后，我们要特别感谢家人和朋友对我们一直的鼓励、耐心和理解。

张彦 Simula研究所，挪威

罗伦斯.T. 杨（Laurence T. Yang）圣弗朗西斯泽维尔大学，加拿大

陈积明 浙江大学，中国

目　　录

第1部分　RFID

第2部分 WSN

第3部分　RFID 与 WSN 集成

第1部分 RFID

第1章 RFID的媒体访问控制协议

由于射频识别（Radio Frequency Identification，RFID）系统不需要视距通信，能够承受恶劣的物理环境，并能够保证低成本和高效能的操作，允许多标签的同时识别，因此，它克服了包括条形码系统（barcode system）、光学字符识别系统（optical character recognition system）、智能卡以及生物测定学（biometrics）（声音、指纹和视网膜识别）等在内的识别系统的不足。

与其他射频系统一样，RFID系统也需要媒体访问控制（Medium Access Control，MAC）协议来防止不同类型的碰撞，因为这些碰撞会浪费网络资源，并减缓识读的过程。由于RFID标签有限的处理和存储能力，在RFID系统中，MAC协议的效率也是重要的考虑因素。

本章主要给出RFID系统中最新的MAC过程的概述。本章中，我们强调信号、存储和处理等不同过程中的需求，根据名义上的价值区分不同的方案。进一步识别实际环境中MAC协议的工作机制，例如在移动的、高密度的阅读器或者是高密度标签等情况下如何影响整个RFID系统的性能。最后，本章将会关注RFID与无线传感器网络（Wireless Sensor Network，WSN）潜在的一些集成。

1.1 概述

射频识别（RFID）技术是一种重要的新兴的自动识别技术。由于RFID不需要视距通信，能够承受严酷的物理环境，并能保证低成本和高效能操作以及多标签的同时识别能力。因此它具有条形码技术、光学字符识别系统、智能卡和生物测定学（声音、指纹和视网膜识别）等其他识别系统所没有的优点。RFID可以很容易地把各种物品的信息发送到移动网络节点，这些移动网络节点可以跟踪、监视，并能够触发动作，或者是响应动作请求。

一个典型的RFID系统由应用主机、RFID阅读器以及一系列的标签组成。标签用来存储一定的信息，标签的存储容量一般在32bit～32000B。标签可以分为主动标签和被动标签。被动标签没有直接的电源供应，它通过接收阅读器发射的电磁

波来产生能量，以便于供给足够满足处理和通信时对于能量的需要。被动标签是目前RFID市场上最常见的一类标签，由于没有电源，限制了它的处理和通信能力。所以，被动标签只能处理简单的状态工作，没有媒介监听能力。而主动标签具有直接的电源供给，可以满足标签处理时对于能量的需要，并且主动标签具有可以处理某些如温度或是大气压数据的能力。RFID阅读器则是标签的汇聚部分，同时也是应用主机的下一级。图1-1描述了这种主从结构。

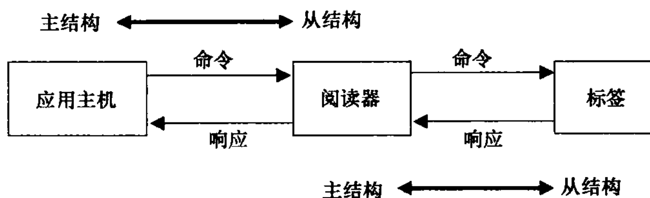


图1-1　RFID系统的主从结构

当某个RFID阅读器定位到一个标签，即表明这个标签在此阅读器的识别区内。一般地，这个识别区是由阅读器产生的电磁波所覆盖的一个物理区域。在这个区域中，阅读器可以给标签提供能量，接收标签的信号，并能对标签进行解码操作[这个过程常被称为独占（singulation）]。独占时，在一个阅读器的识别区中，阅读器只能识别一个标签，而一个标签也只能被一个阅读器识读。一个阅读器识别区中的标签和具有重叠识别区的阅读器可以同时接入无线媒体来进行数据通信。

但是，这种同时的无线媒体接入将会导致数据碰撞，并影响整个RFID系统的性能。为了保持系统的可操作性，有效的媒体访问控制（MAC）机制是非常有必要的。与其他射频系统一样，MAC机制的主要目的在于通过一种主动的或是被动的方式来调节对媒介的访问，以达到降低碰撞的效果。主动时，有效地分发关于共享媒体访问请求的信息来避免碰撞；被动机制是响应碰撞，并试图加快系统碰撞后的恢复过程。传统避免碰撞的方法，例如载波监听多路访问（Carrier Sense Multiple Access，CSMA），其并不适合RFID系统，尤其是当被动标签的能量受限，并且是基于反射[即使用反向散射调制（backscattering modulation）]来进行通信的情况时。这种避免机制同时提高了整个标签的成本，缩小了单个阅读器的潜在识别区域。因此，RFID系统更倾向于采用被动的机制来减缓碰撞。在RFID系统中，碰撞可以根据实体的类型，分为以下几类。

1）多标签与单阅读器的碰撞：如图1-2所示，此类碰撞主要发生在一个阅读器的识别区内，一个以上的标签试图同时响应阅读器的请求时。多标签与单阅读器间的碰撞危害极大，尤其是在使用被动标签的环境中。此类碰撞会导致识读速率降低、资源浪费、延时增大等后果。

2）多阅读器与单标签的碰撞：如图1-3所示，此类碰撞主要发生在单个标签

被多个阅读器识别的情况下。这种情况下，多个阅读器试图独占单个标签，将会导致标签间隔状况出错。结果是，这个标签将不会被检测到。

3）阅读器与阅读器间的碰撞：此类碰撞是由传统的频率干扰引起的，具有相互干扰区域的多个阅读器会被同一频率锁定。现存的跳频、动态频率分配以及动态能源调节等机制可以隐藏这些碰撞。

本章的目的是整理各类文献已报道过的应用于 RFID 系统的防止碰撞的机制，并在本章中提供与其他文献相比较的、更加综合的，并且最新的 MAC 协议[1,2]。本章最后总结了 RFID 系统中与 MAC 协议相关的各个主要领域。简要阐述未来可能流行的基于 TDMA 解决方案的原因。这部分综述主要放在了 1.2 节。在 1.3 节将主要阐述标签碰撞，以及基于标签需要的存储和处理的不同解决方案的分类。1.4 节总结解决阅读器碰撞的相关建议。在 1.3 节和 1.4 节中，我们将会比较基于通用解决需求的不同方案。最后在 1.5 节中，通过阐述当前和将来在 RFID 系统的 MAC 协议中将遇到的机遇与挑战总结整个章节。

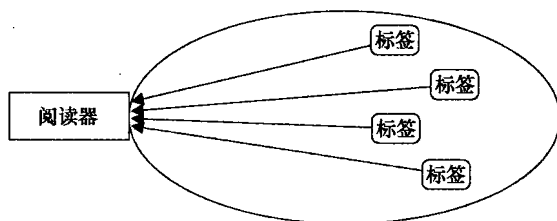


图 1-2　多标签与单阅读器的碰撞

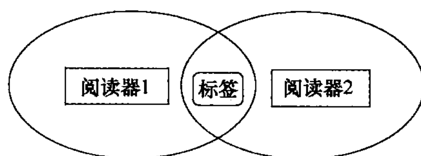


图 1-3　多阅读器与单标签的碰撞

1.2　RFID 系统 MAC 协议的预备知识

在射频系统中，MAC 协议提供了信道访问控制机制，允许多个设备共享同一物理媒体。

其中，最著名的 MAC 技术基于 CSMA 的 MAC 协议、多路访问碰撞避免（Multiple Access With Collision Avoidance，MACA）以及 Aloha 协议。RFID 系统在阅读器与标签的通信中使用半双工的，点对点的连接，并且阅读器与标签的通信是基于反向散射调制的。在反向散射调制中，标签利用反射信号，把串行数据通过天线发送出去。由于反射通信的特性，使得传统的 MAC 技术并不能使用在标签上，因为这些标签不能监听媒体，监测碰撞或者监听到其他信道阻塞的存在。这意味着所有避免碰撞的情况和解决机制都无法在标签上实现。因此，RFID 系统中的 MAC 协议只限于在解决阅读器的数据碰撞上，这样就产生了在标签碰撞和阅读器碰撞中都适用的防碰撞算法，这些算法主要使用基于 Aloha 协议的竞争避免方案。

RFID 系统的运行状况是：在短时间具有非常高的活跃期，而其他大量的时间都处于相对静止的不活跃状态。对于一个阅读器来说，在它识别区域内的标签数量通常是不确定的。这个问题导致了必须设计具有可伸缩性的需求和具有适应性的协议。通常地，有 4 种方法可以用来解决无线技术中的多路访问问题。这些方法基于时间、空间（位置）、频率和码制来调节媒介的访问。码分多址（Code Division Multiple Access，CDMA）使用扩频技术来支持高速率数据复用。在传统的扩频技术中，每个用户使用正交码来编码数据包，从而使多个用户能够同时传输数据。但是，复杂的接收机设计以及高的计算和能量需求限制了 CDMA 技术在 RFID 系统中的使用。最近参考文献［3］报道了基于结合 TDMA 和 CDMA 技术的方案。但是对 CDMA 方案在 RFID 系统中的使用还没有进行深入的研究。

空分多址（Space Division Multiple Access，SDMA）技术[4,5]在空间上复用信道。空分多址运行过程的原理是，在同一时间，同一地点，只存在有限数量的 RFID 标签，独立于整个识别区域可能出现的标签总量。因此，在空间上可以孤立这些 RFID 标签，使得由其他标签造成的干扰降到最低。通过调节阅读器的能量（能量控制[6-8]），使用适应性阵列（adaptive array）、多入多出（Multiple Input Multiple Output，MIMO）天线技术[9]，以及电子控制的定向天线[5]技术，可以形成许多个空间孤立区。

通过调节阅读器的发射功率来改变阅读器的识别距离，基于簇的能量控制算法[6,8]可以把识别区分成更小的簇。每个簇内的标签被分别识读。由于一个簇内标签的数量明显少于整个识别区域内标签的数量，碰撞将随之大大减少。如图 1-4 所示为一个识别区域分割的例子。图中，识别区域被分割成 3 个簇：d、d′和 d″。当阅读器发送一个请求信号时，只有在当前簇的标签才会响应。比如，假设阅读器已经识读 d 簇内的标签并又把请求发送到 d′簇，在这个例子中，只有 d′簇内的标签，即标记为 T′的标签会对请求作出响应。标记为 T′的标签都被识读之后，它们会进入休眠模式。然后识别区域将会增大到 d″簇，并发送新的请求，以此类推。同样的传输控制方案已经运用到了避免阅读器碰撞上[7]，在这种情况中，阅读器的功率将会被调节，以减小识别区域的重叠。

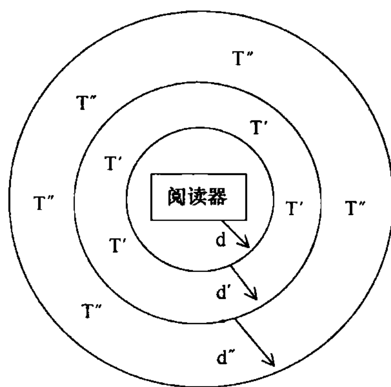


图 1-4　阅读器和标签之间基于距离分簇，使用阅读器的功率级别来调整它的广播域

通过应用例如自适应阵列天线、多入多出（MIMO）天线[9]等的智能天线技术，可以降低碰撞，提高系统的吞吐量。自适应阵列天线通过其他天线单元来抵消

干扰，即其中一个天线接收的信号可以被其余天线利用来去除干扰，以最大化接收信号的强度。在多入多出（MIMO）天线技术中，每个天线接收一个重叠的具有不同空间信号的多路传输流。在接收机端，可以使用这些不同点，通过信号处理来分离复用的数据流。这两项技术可以减少碰撞的产生，但是阅读器的成本将会显著地增加。

通过使用电子控制定向天线[5]，可以适应性地调节 RFID 阅读器的定向波束，以使阅读器一次识别一个标签子集。这种技术通常被称为自适应的空分复用技术，如图 1-5 所示。为了识读某一特定的标签或是识读识别区域内的所有标签，阅读器使用定向波束扫描整个识别区域，直到目标标签或是所有的标签被识读为止。空分多址（SDMA）技术可以有效地降低碰撞，但是也需要付出较高的成本才能实现复杂的天线系统，因此这也限制了 SDMA 技术的一些具体应用。一种最可行的方法是前面提到的能量控制方案[6]，它的成本较低，并且能够有效降低阅读器和标签的碰撞。

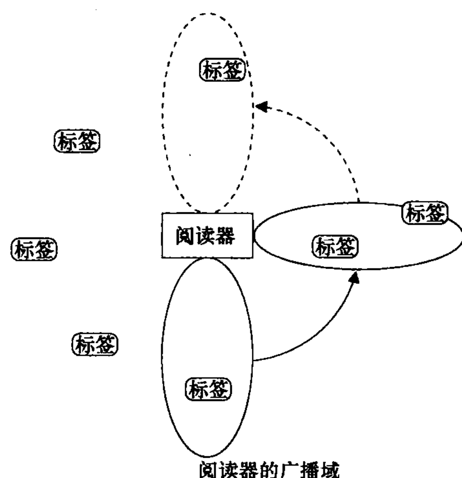


图 1-5　带有电子控制定向天线的自适应 SDMA

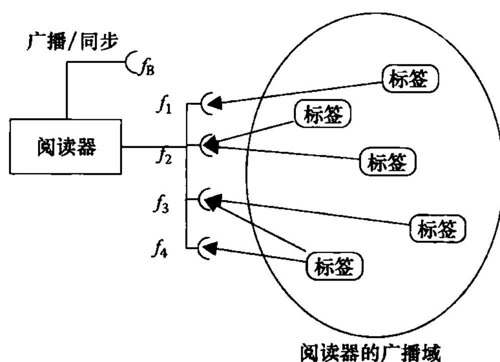


图 1-6　FDMA 程序，几个可用的用于标签和阅读器之间通信的频道

频分多址（FDMA）技术使用多个信道，每个信道使用不同的载波频率进行通信。在 RFID 系统中，这可以实现阅读器广播频率和多个标签的多个频率进行锁定的可能性。阅读器使用广播频率来同步和发布识别命令。标签使用许多可用频率中的一个，如图 1-6 所示的 f_1、f_2、f_3和f_4来响应阅读器。频分多址的优点是在多标签和多阅读器同时通信时，对非干扰频率的利用。但是，频分多址技术并没有在 RFID 系统中得到广泛的应用，其原因是 FDMA 技术对于标签来说并不实际，因为这样阅读器需要相对较高的成本，而且每一个接收机也必须提供独自的接收信道，这就限制了频分多址技术在特定场合的应用。

时分多址（TDMA）技术把可以利用的信道在潜在的参与者之间根据时间维度

进行划分。例如，在GSM、蓝牙和IEEE802.16（Wimax）等这些主要的网络中，时分多址技术通常是采用改进的，并且大多数情况下是混合的方式进行使用的。到目前为止，在RFID系统中，时分多址技术是最主要的媒体访问协议。这是因为与FDMA、CDMA和OFDMA等技术相比，TDMA技术简单，且相对于被动标签来说，具有较低的处理成本，以及在计算、处理和资金消耗方面也不复杂。如图1-7所示，TDMA的处理过程进一步可以分为标签驱动和阅读器驱动两类。标签驱动过程采用异步的方式，阅读器不控制数据的传输。标签驱动过程运行较慢，而且不灵活，从而限制了它的应用。于是，大多数TDMA过程使用阅读器驱动的方式。阅读器驱动方式采用同步机制，阅读器操作所有标签的数据传输，即控制标签何时传输数据，选择哪个标签传输，在给定的时间内传输数据。这种选择过程可以从一组标签中选择某一特定的标签，这称为独占操作。阅读器驱动的方法用来处理RFID系统中的数据碰撞问题。

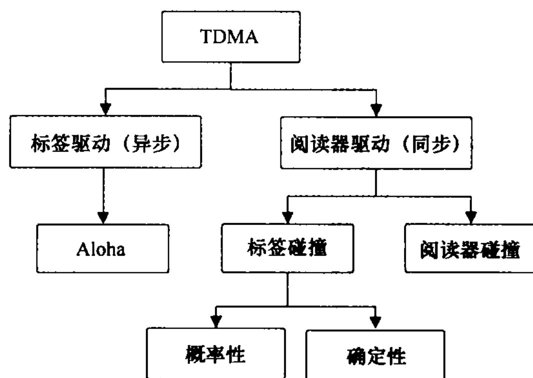


图1-7 TDMA程序的分类

标签碰撞机制经常使用基于树算法或者是概率帧Aloha方案的标签数量的逻辑划分，把它转化为可以更方便处理的标签集。阅读器与阅读器之间的碰撞、阅读器与标签之间的碰撞使用诸如时间安排、干扰学习、色彩方案等传统竞争解决方法来处理。标签与阅读器之间的碰撞的各种解决方案是根据它们使用的技术来分类的，接下来的几节将按顺序介绍各种方法。

1.3 标签碰撞

在RFID系统中，最常见的碰撞来源即是标签碰撞。标签碰撞主要发生在多个标签同时在一个阅读器的识别区域内，并且同时响应这个阅读器的请求命令时。由于被动标签没有载波监听或是标签间通信的能力，所以它被设计成阅读器驱动的方式。因此，阅读器使用采集技术来解决标签碰撞，这种采集技术被称为防碰撞方

案。如图 1-7 所示，根据相关文献，这些防碰撞方案被分类成确定性机制或是概率性机制[1,5]。

在确定性机制中，阅读器在给定的时间内，分裂和识读一个标签集合。分裂是基于先前过程循环所获得的竞争信息，以及试图减少下一个循环的竞争信息。根据它们的分裂方法，确定性的防碰撞机制是根据基于树的算法进行分类的。确定性机制使用标签的序列号（识别码）或者随机产生的数字来实现树枝的分裂。在既定的环境中，确定性方法可能需要花费相当长的时间。但是，这并不会造成标签的饥饿问题。在饥饿状况时，标签可能在长时间内不能被识读；更有甚者，这个标签可能根本不能被识读。

在概率性机制中，阅读器通过特定长度的帧进行通信，而标签随机地在帧中发送一个特定的时隙。帧的大小根据先前交互识别循环（Interrogation Cycle）过程所得到的信息进行调整，并根据标签的密度和分布来提高适应性。帧的处理过程一直重复进行，直到所有的标签都被识读为止。由于较低的消耗，所以概率性方法的执行过程十分迅速，但是会受到标签饥饿状况的影响。

1.3.1 确定性的防碰撞机制

确定性的防碰撞机制本质上是基于树的防碰撞算法。基于树的算法是一种两阶段握手算法，包括阅读器与标签之间的一系列称为交互识别循环的交互。这种交互识别循环的目标是使用序列号（ID）或者随机产生的数字将许多标签分裂成可以更方便处理的标签集。基于位碰撞和它们各自的位置，可以把二进制树分裂成两个树枝。这个位置是由先前的交互识别循环得到的。要想获得比特级的碰撞信息就需要由阅读器识别的碰撞的精确比特位置。出于这样的原因，一般地，RFID 阅读器需要曼彻斯特编码或者是不归零码（non-return-to-zero，NRZ）比特编码。

为了更好地理解基于树的确定性算法，我们将跟踪传统的二进制搜索树算法[10]的执行。使用多个交互识别循环的二进制搜索树算法的目的是从一个更大的标签集中独占（singulate）一个标签。在每个交互识别循环期间的两次握手是阅读器为了接下来的交互而基于先前交互来广播的命令。这个命令集由以下四个主要的命令组成：

1）请求：携带一个作为标签参数的序列号。如果一个标签自身的序列号小于或者等于接收到的序列号，标签即会把自身的序列号发送回阅读器，否则，标签就不会响应。

2）选择：携带一个作为标签参数的序列号。只要具有特定序列号的标签才会被选中，来处理诸如读写数据的其他命令。只有被选中的标签，才会继续响应阅读器的命令。

3）读数据：被选中的标签将存储的数据发送给阅读器。

4）取消选择：取消选中标签的活动状态，使标签休眠。此外，直到标签被阅

读器复位之前，非选中的标签将处于非活动状态，不会响应阅读器的进一步请求信息。

假设在阅读器的识别区有三个 RFID 标签，并且三个标签都具有四位的二进制 ID 号，分别是 1010、1011 和 1110。独占操作从阅读器发送碰撞概率最高的序列号开始。目标是为了使所有的标签能得到响应，以便于检测所有标签 ID 之间的位碰撞。

如图 1-8 所示的例子中，算法的交互过程由阅读器发送序列号为 1111 的请求命令开始。1111 是在这种情况下碰撞概率最高的序列号。由于在识别区内标签的序列号都小于请求的序列号，所以三个标签都将回复它们的 ID 号。此次交互过程导致了第 0 位和第二位数据的碰撞（Collision，C），即最小意义位（Least Significant Bit，LSB），此时为 1C1C。

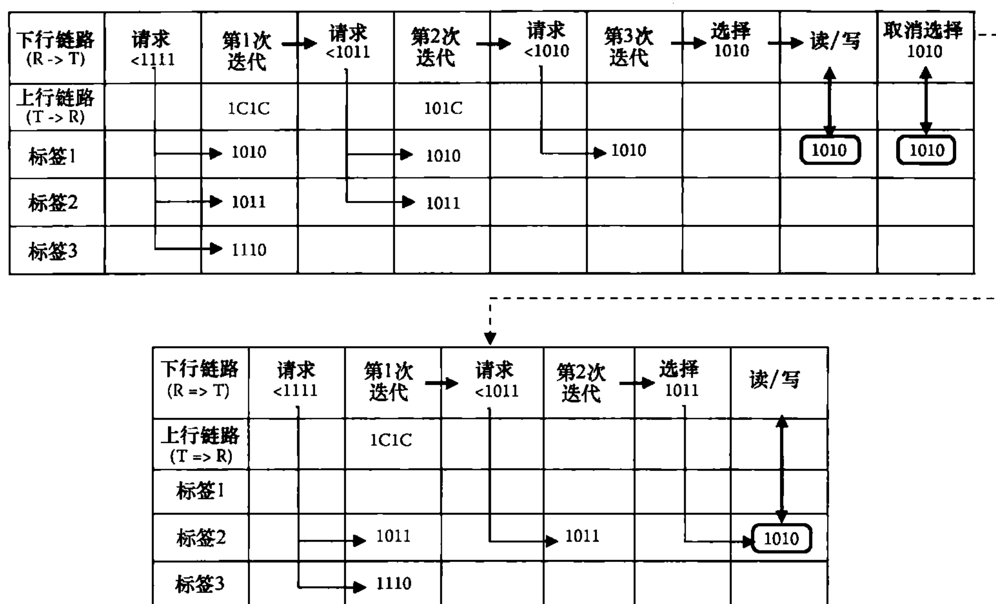


图 1-8　使用三个标签和单一阅读器的二叉搜索树防碰撞算法的执行路径

在第一次交互中，第三位是碰撞发生时最有价值的一位。这意味着至少存在序列号为 1100 和 1011 的两个标签中的一个标签。此分裂中，二进制搜索算法把搜索分裂成两个子集，以试图限制子交互的搜索区域。此算法把第三位设置为 0，即 1011。为了试图捕捉到所有的两个最大意义位（MSB）是 10 的标签，第三位之后的 LSB 都被设置成了 1。阅读器开始广播 1011 的请求命令，标签 1 和标签 2 符合此命令的标准。响应的这两个标签又会造成在第 0 位的碰撞，即 101C。阅读器重复这样的分裂过程，并在随后的交互中选择 1010 作为请求串。只有标签 1 符合此

请求的标准，从而实现了对标签 1 的独占操作。现在，不需要进一步的交互，阅读器已经可以在没有碰撞的情况下成功地监测到了单个标签。使用随后的选择命令，使用检测到的标签 ID 地址选中标签 1，并可以在没有其他标签干扰的情况下，进行数据的读/写操作。此时，在选择性的读数据命令中，其他标签是沉默的。图 1-9 所示为请求响应的树。尽管每个相邻循环的字符串是接收的响应，树节点上显示的字符串是阅读器请求和发送的。一个节点分裂成两个子节点，并在 MSB 碰撞中，分别在子节点的左边和右边，增加 0 和 1。这个过程将持续到 1010 被独占为止，在图 1-9 中以双同心圆区别开来。

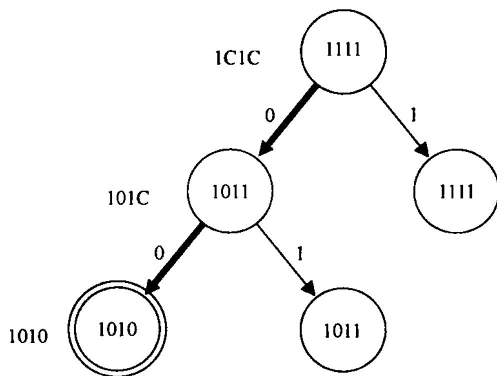


图 1-9　样本 RFID 系统的分离树

在完成需要的读写过程之后，使用取消选择命令将选择的标签（标签 1）休眠。休眠的标签处于完全的非活动状态，不会响应接下来的任何请求命令。休眠的标签减少了接下来独占操作的交互过程中响应标签的数量，从而降低了碰撞的产生，以及减少了交互的次数。参照图 1-8，标签 2 的独占操作将比标签 1 的独占操作少一次交互过程。标签 1 的休眠对于减少碰撞的发生起到了积极的作用；因为如果标签 1 没有休眠，它将会再次与标签 2 碰撞。

如图 1-10 所示，根据是否使用碰撞跟踪或是碰撞监测方法，将确定性的防碰撞机制进一步进行分类。

（1）碰撞跟踪　在碰撞跟踪方法中，阅读器和标签都保存了前一次交互识别循环的一定数量的碰撞信息。这其中大部分指出了近期的查询、碰撞位，或者是树节点的信息，被用来发送下一个识别循环的查询。二进制搜索算法和它的变种[11-22]，归于确定性防碰撞算法的碰撞跟踪类。传统二进制树算法变种的主要区别在于它们增强了目的性，即缩短了执行的时间，减少了存储的使用，消除了不必要的交互过程。

在传统二进制树算法中，当标签被成功的独占之后，进行读写和休眠，为了随后的标签初始化，都需要从一连串节点的根节点进行。例如，在上面的例子中，当阅读器发送 1111 的初始查询时，即最大可能的序列号，标签 1 将被独占以及取消选择。这将导致通信和处理过程中不必要的重复。不过这种重复在对标签进行成功的独占之后，通过设置顺序的查询点，是可以消除的。这可以解释为选择当前查询节点的父节点或是它的兄弟节点[20,21]。这种简单的修改，加速了交互识别过程，减少了重复保存的信息。传统的二进制基于树的算法在识别过程中，通过设置最大

意义碰撞位为 0 或者 1，一次只能识别一个标签。参考文献 [14，17] 指出这样会导致较低的性能，并认为在同一时间，树导航处理两个标签时，没有任何基础性的改变。这发生在顺序碰撞的最后一位，即当 LSB 碰撞解决的时候。当 LSB 碰撞发生时，很明显预示着有两个标签的存在，即位 0 和位 1 的碰撞，于是加速了识别过程。许多文献已经报道了使用例如干扰记录[18]，实现标签 ID 号之间的相似化[23]，以及使用支持标签状态的高度导向，标签状态支持的深度优先的搜索算法[24]，和固定变量的树算法[12,22]。所有的这些方法在没有显著限制传统算法的基础上增强了二进制树算法。

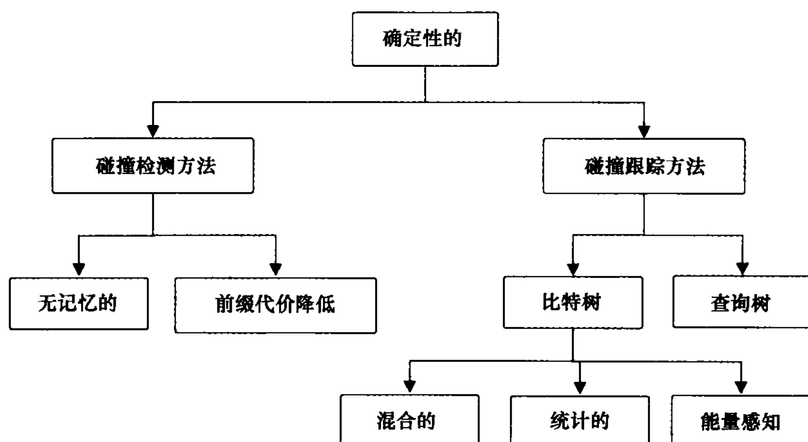


图 1-10　标签碰撞的确定性 TDMA 模式的分类

如图 1-10 所示，可以进一步把跟踪碰撞方法分为统计的、混合的以及具有能量感知的。静止方法获得存储在先前交互识别循环中的信息，以协助接下来的循环。适应性的二进制分裂算法[25-28]通过使用在最后一次识读循环中获得的信息初始化传统的树分裂算法。先前循环识别的固定标签（即静止标签），将会在下一次识别循环中被阅读器再次识读。由于阅读器已经知道了固定标签，所以可以避免碰撞的发生，而传统的方法用来识别移动标签。移动标签被定义为从外界进入的标签，因此这些标签是没有在最后一次识别循环中出现的，或者是从阅读器的识别区域移出的标签。在需要支持移动，但是又需要大量存储空间的情况下，这些适应性的分裂算法可以显著地减少碰撞的次数。由于很大一部分标签是在移动的情况中，所以适应性分裂算法的性能并不比传统的二进制算法要好。

混合树方法[29]是结合了确定性方法和可能性方法。它结合了反馈隙（slotted back-off）机制的基于树的协议。此算法使用四进制（4-ary）树，替代了传统的二进制（binary）树，而反馈隙的目的是减少不需要的以及不必要的查询命令。当响应阅读器时，标签使用它一部分的 ID 号来设置反馈定时器。如果有碰撞，阅读器能够部分推测到标签的 ID 号是怎样被分发的，并潜在的降低不需要的循环。正如

参考文献［26］报道的静止方法，这种方法可以得到前次识别循环的信息来避免静止标签的碰撞，并且从根节点到适应标签的移动性来开始查询过程。另外，为了克服静止方法的不足，混合方法如反馈定时器等，都要求额外的功能，这增加了标签的成本。

为了减少碰撞，MAC 进程以增加能量消耗为代价。这种能量消耗主要是由于处理 MAC 协议而产生的额外的处理和通信。当保持低碰撞时，能量感知方法[15,30-32]主要是用来处理包括标签能量消耗在内的能量问题。参考文献［15］中的方法实现了传统的二进制树算法，并在阅读器检测到位碰撞时，能够采取相应的动作。这可以使阅读器把特殊标记发送给标签来停止标签发回数据。于是，标签会发送更少的数据，从而减少通信的过程，间接地减少了能量的消耗。更多有效的方案[32]是在每个树节点上使用多时隙，以及三个不同的防碰撞协议。动机是为了监测标签，因为二进制树算法发送大量的查询给标签时，会产生很多碰撞。这在传统的二进制树算法中是很重要的，因为它将决定在接下来的识别循环中去查询哪个子节点。标签被允许在一个时隙帧中发送响应，以便于避免碰撞。这种查询可以用来发现碰撞子树，以及识读标签。这两种机制使得通过更少的查询就可以来读更多的标签，因此减少了标签上的能量消耗。

（2）碰撞检测　碰撞检测方法是无状态的算法，所以它们不用学习跟踪。与碰撞跟踪方法相关，检测算法在识别区内可能需要花更多的时间来识读所有的标签。它们的主要优点是它们不需要存储能力，特别是对于标签和它们相对低的通信能力来说。如图 1-10 所示，可以把碰撞检测方法分为无存储（memoryless）的方法和前缀减少（prefix reduction）的方法。

查询树协议（QT）[33]以及它的其他变体[15,28,34]是无存储方法的主要例子。在这个方法中，标签除了存储其自身的序列号之外，不需要额外的存储空间。这意味着不需要存储空间来存储在二进制碰撞检测跟踪树算法中的随机序列号、指针或是状态。查询树协议使用标签的序列号来处理树分裂机制。阅读器使用一串位变量作为参数来广播查询。标签收到查询，并用广播的序列号与它自身的序列号的 MSB 来进行匹配。如果匹配，则标签发送其剩下的标签号的最小显著部分（least significant portion）；如果不匹配，标签将会保持沉默。阅读器保持发送位字符串的队列。在帧的开始处，队列使用两个 1 位的字符串 0 和 1 来初始化。阅读器提取队列中的一个位串，并广播给它的识别区中的所有标签。如果标签发生碰撞，阅读器将把与最后发送的位字串相比的两位更长的位字串压入队列。扩大查询，标签位字符串遍历从 MSB 到 LSB 的可能性序列号，直到所有的标签最终都能得到响应为止。与二进制树相反，查询树使用简单的功能，并且不需要标签提供的状态和指针。但是，标签序列号的数据分发将显著影响查询树协议的读标签延时。当标签具有相似序列号时，将会增加标签的读延时。

当位的数量作为查询参数发送时，查询树和它的变体是有用的。原因是从最大

显著端（most significant）传输到最小显著端（least significant），位字符串会与序列号的长度相等。在识别循环期间，由于大量的位传输，标签和阅读器将同时会出现通信的开销。

前缀开销减少（prefix overhead reduction）算法[16,18,19,35]保持前缀和交互开销减少的方法，增强了无存储机制的性能。前缀随机查询树（Prefix Randomized Query Tree，PRQT）[18]基于前缀减少机制，通过使用根据标签长度和它创造的更长的读循环克服了基于查询树算法的限制。前缀随机查询树与查询树相似，是一种无状态的方法，但是与查询树算法不同，前缀随机查询树选择标签随机产生的前缀，而不是使用基于ID的前缀，因此需要提供额外的存储空间。每个标签产生一个随机长度的前缀，这个随机前缀将在识别过程中使用。理想长度的前缀取决于对识别区内标签数量的估计。标签估计是由适应性的理想增量前缀长度算法[35]提出的。算法一开始设置一个小的初始前缀长度，并轮询所有可能的前缀。初始前缀长度随后增加，直到碰撞比例满足预先设置的状态为止。为了减少前缀开销，碰撞检测机制需要最小限度的额外存储空间来满足在标签端有快速的读速率和较低的开销。

1.3.2　概率性的防碰撞机制

在概率性机制中，阅读器用帧长度通信，标签用帧中一个特定时隙来传输数据。阅读器重复这个过程，直到所有的标签成功地传输一次数据为止。阅读器控制的同步是非常有必要的，因为标签不得不在它的时隙帧中传输数据。帧大小可以根据来自先前识别循环的碰撞、空闲和占据的帧信息进行调节。根据标签的密度和分布可以提高帧的适应性，因此减少了空闲帧和碰撞帧。由于概率性方法具有更低的开销，所以与确定性方法相比，它的速度更快。但是，概率性方法存在着标签饥饿问题。

可以使用参考文献［5］的例子来说明1.3.1节中传统时隙Aloha防碰撞过程的细节。假设仍有ID号分别为1010、1011和1110的三个标签。与确定性方法相似，概率性方法也使用一串对话协议的交互命令进行通信。命令的设置如下。

1）请求（REQUEST）：使用一个时隙，同步和提示所有的标签发送它们的序列号给阅读器。在举例的RFID系统中，存在着三个有效的时隙。

2）选择（SELECT）：发送一个作为参数的具体序列号给标签。只有那些参数与序列号匹配的标签被标记允许用来进行进一步的读写操作。但是，与序列号冲突的标签，仍旧会响应请求命令。

3）读数据（READ_DATA）：选择的标签将它存储的数据发送给阅读器。在实际的RFID系统中，其他一些命令也是有效的，但是为了简化，在这里被忽略了。

图1-11所示为时隙Aloha协议的执行轨迹。在一定的时间间隔之后，阅读器周期性地发送请求命令。在接收到请求命令之后，标签随机地选择三个有效时隙中的一个。标签利用时隙把它的序列号发送给阅读器。由于随机时隙的选择方法，在

时隙 1 中，标签 1 和标签 2 会发生碰撞，只有标签 3 可以顺利把它的序列号发送给阅读器。顺利发送序列号的标签将会使用选择命令，在被选中后进行接下来的读写操作。如果没有发送是成功的，那么请求命令将会重复地发送，直到成功地接收到标签的序列号，以及在帧中没有检测到碰撞为止。正如观察到的，在概率性碰撞算法中，标签在帧中随机选择一个时隙，并使用它选择的时隙来响应阅读器。

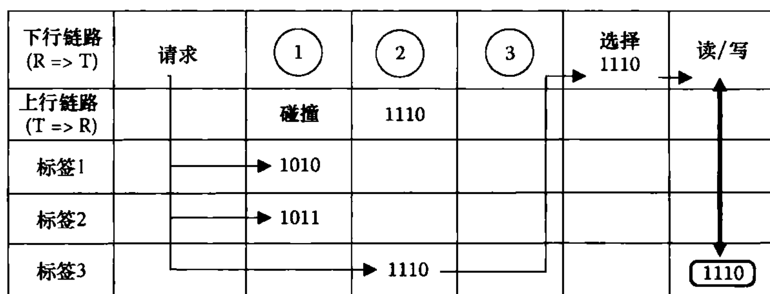


图 1-11　传统时隙 Aloha 协议执行轨迹，使用三个标签和一个单一阅读器的样本 RFID 系统

当标签的数量较少时，标签碰撞的概率会很低，用来识别所有标签的时间也相对较短。

但是，随着标签数量的增加，碰撞的可能性将随之增大，识别标签需要的时间也会显著地增加。因此，概率性算法的性能取决于在阅读器识别区中标签的数量。如图 1-12 所示，一般可以把概率性机制分成两种：静止的和动态的。静止方法中时隙的数量是固定的，大部分是在标签密度较低的情况下使用。静止算法的例子有基于 Aloha 和帧隙 Aloha 的算法。一个固定帧大小可以使得实现较为简单，但静止算法也造成了较低的效率。例如，在标签数量较少时使用长的帧，或者是标签数量多时，使用较短的帧，都会导致延迟以及资源没有有效的利用。这些问题在动态方法中可以得到解决，动态方法中帧的大小可以根据标签在识别区中碰撞的概率进行调节。例如，动态帧隙 Aloha（Dynamic Frame Slatted Aloha，DFSA）算法[5]基于诸如识别标签的时隙数量和碰撞的时隙数量来调节帧的大小。

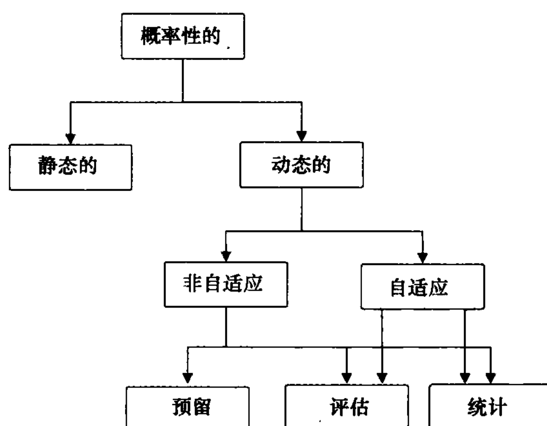


图 1-12　标签防碰撞的概率性基于 TDMA 的模式的分类

不同的 DFSA 算法的变体改变帧大小的方法不尽相同。其中一个变体中，帧大

小是基于事先定义的一些阈值来进行调节的。例如，当碰撞隙数量超过预先设置的值时，帧的大小将会增大。当碰撞减少时，帧的大小也会降低到先前的值。这种帧大小的动态调节允许阅读器根据标签的密度来改变它的帧大小。换句话说，帧的大小会随着碰撞增加而提高，或是随着碰撞的减少而减小。

在另一个变体中，识别循环以设置初始帧大小为2或4开始。随着不成功传输的增加，帧的大小会增加，直到一个标签顺利地传输为止。如果至少一个标签可以被成功地识别，当前的读过程将会被取消，并重新初始化为初期的帧大小。尽管DFSA和其他变体基本具有所有性能，但是改变帧大小仅根据标签的数量是不够的，因为帧的大小不能无止境地增加。例如，上述的第二个变体，当标签的数据较少时，阅读器可以在没有过多碰撞的情况下识别所有的标签。但是，如果标签的数量很大时，而阅读器又总是以初始的最小帧大小开始，所以时隙的数量需要以指数的形式增加。

如图1-12所示，可以进一步把动态算法分成自适应和非自适应两类。自适应算法使用统计性的信息来调节帧的大小。自适应算法比非自适应更适合标签移动。因为当同时识别RFID标签时，自适应算法可以减少标签碰撞的概率。动态算法包括自适应和非自适应的，可以根据保守，估计和统计技术的使用进行进一步的分类。基于估计的方法是最重要的基于Aloha的方案[10,37-43]，可以根据相应的标签数量估计技术来调节帧的大小。优化的帧隙Aloha（Advanced Framed Slotted Aloha, AFSA）算法[10]在初始化读过程之前，先估计标签的数量，并基于这个估计调节帧的大小。契比雪夫（Chebyshev）不等式估计函数用来估计标签的数量。但是，AF-SA算法增加帧的大小是无止境的，这在标签的密度较高时显得不实际。

增强的优化动态帧隙Aloha算法（Enhanced DFSA, EDFSA）[42]通过对最大帧大小设置边界估计（即设置一个帧大小的上界）来克服AFSA的不足。EDFSA使用AFSA估计函数来初始估计没有被读的标签数量，然后按比例把没有被读的标签分成数个组。只有其中一个组的标签允许在某一个时间同时被响应。接下来读循环的最大吞吐量的分组数量在每次识别循环后被重新计算和调整。当阅读器广播一个请求时，它会把标签分组的数量和一个随机号发送给标签。接收请求的标签根据接收到的随机号和标签自己的序列号产生一个新的号，并根据标签分组数分解新号。只有具有零的余数标签才会响应请求。估计和没有被读的标签的组在每次读循环之后重新被执行，直到所有的标签被读过为止。尽管EDFSA克服了AFSA协议的不足，但是与AFSA相比，它的读循环时间更长，在标签上也需要额外的功能。

统计性的算法[16,41,44,45]使用统计性的信息改善RFID系统的读标签时间。适应性的时隙Aloha协议（Adaptive Slotted Aloha Protocol, ASAP）[41]利用从先前识别循环和读过程中获得的标签数量相关的信息，来估计当前阅读器识别区域中标签的数量。基于估计算法的最大似然性（Maximum Likelihood, ML）用来实现这个目的。调节最佳的帧大小来体现标签的估计。当在每次识别循环的开始初始化估计和帧调

节时，通过计算标签到达和离开的速率来支持标签的移动性。统计性算法类似于确定性碰撞分类，并共享相同的接口和交换。

1.3.3 讨论

本章到目前为止讨论了大部分防碰撞协议，通过付出其他性能上的代价来获得最大的优势。例如，在最小化碰撞中，算法将会增大读数据的速率。但是，这种速率的提高需要额外的存储器，同时通信和处理的要求增加了标签的成本。接下来，将讨论在设计防碰撞算法时，一些性能上的权衡。

1）速度：读标签时可以达到的速率。这是许多文献理论中所追求的目的。

2）消耗：标签和阅读器的通信和处理消耗。通信消耗包括发送额外的位，这些位本是不必发送的。处理消耗包括除了一般之外的额外识别循环或是数据运算。

3）状态：状态的数量可以有效地存储在标签内。额外的存储意味着需要额外的存储空间。阅读器可能保持统计性的信息，但在阅读器上存储的成本明显少于标签上存储器的成本。

4）移动性：在询问过程期间，能够适应标签进入和离开识别区域的能力。

5）规模：适应标签高密度部署的能力。

6）成本：在标签和阅读器上需要额外的功能和处理能力。增加成本同样意味着需要增加对设备金钱上面的投入。

在确定性和概率性算法中，分别用“√”描述了方案的改良点，用“↓”描述权衡，见表 1-1 和表 1-2。方案可能不会绝对地影响性能，但可以间接地影响，这以“—”来描述。例如，前缀减少方案主要是为了减少通信的开支，这可以间接地减少在识别标签时的时间花费，因此可以提高标签的识读速率。这在表中已得到了指出，确定性的或是概率性的方案在不同的环境下，伴随着不同的权衡和适应性。因此，不同的标签应用集中，没有一个单一的方案是可以满足全部性能要求的。例如，在一个典型的仓库环境中，阅读器部署在对接门或是传送带处来识别数以百计的移动标签或者是数以千计的物品级别的标签。因为速度、移动性和规模这些因素比降低成本更重要，以一个较好的性能识别所有的标签是非常重要的。与其他高速率的阅读机制，例如限制了可伸缩性或者无存储能力的低移动性的混合型方案相比，统计性的方法是最合适的。一个基于 RFID 的访问控制系统不需要可伸缩性和高的读速率，因此任何一般的基于树的算法都是有效的。

表 1-1 确定性防碰撞算法比较

算法	速度	代价	有状态	移动性	可扩展性	成本
能量感知[15,31,32]	√	√	—	↓	↓	√
混合[29]	√	√	↓	√	—	↓
无记忆[15,28,33,34]	√	↓	√	—	—	√

（续）

算法	速度	代价	有状态	移动性	可扩展性	成本
前缀[18,19,35,36,46]	—	√	↓	—	—	—
统计[25-28]	√	—	↓	√	√	↓

表 1-2　概率性防碰撞算法的比较

算法	速度	代价	有状态	可扩展性	成本
静态[8]	↓	√	—	↓	—
预留[47]	√	↓	↓	—	↓
评估[10,37-42]	√	↓	↓	—	—
统计[16,41,44,45]	√	—	↓	√	↓

1.4　阅读器碰撞

具有重叠识别区域的阅读器可能在标签的识别时产生相互干扰。正如与其他射频系统一样，即使识别区域不重叠，一个阅读器的运行也可能会影响到其他阅读器。在这种情况下，都会导致阅读器的碰撞[48,49]。多频率使用可能会产生阅读器碰撞，RFID标签是限制了功能的设备。这些设备不能区分多个阅读器，它们也不能被大量生产并使用多个频率来进行通信。

如图1-13所示，对阅读器碰撞进行分类。采用基于调度（scheduling）、覆盖以及学习等多种机制来解决碰撞。

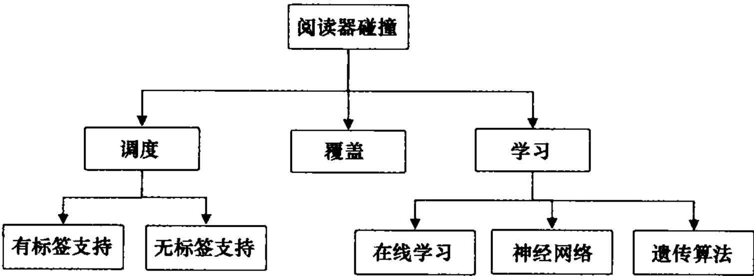


图1-13　阅读器碰撞模式分类

在基于调度的方案中，为阅读器调度了频率和相关的时隙。调度可以是集中式的或是分布式的，并且支持静态的和动态的频率分配。同时，调度需要标签的支持，即标签需要额外的存储空间来存储额外的数据。但是，通常情况下调度不需要标签支持。诸如色彩波形算法[50,51]就是这一类别。另一方面，基于学习的方法是基于等级制度的在线学习、遗传算法和神经网络方法。学习方法通过学习碰撞模式

以及基于学习模式的分配频率来最小化碰撞。例如 HiQ[52]算法就是基于此类方法。已有文献报道了例如使用灯塔（beacon）信道[53]、中心协调器[54]，以及作为覆盖问题处理阅读器碰撞这样的交替方法。

为了优化频率信道的分配，参考文献[50,51]报道一对被称为分布式色彩选择（Distributed Color Selection，DCS）和可变的分布式色彩选择（Variable-Maximum Distributed Color Selection，VDCS）。DCS 和 VDCS 都是基于调度的类别。被称为色彩波形的两种算法的目的是使用多彩色给多阅读器网络上的每个节点上色，这种方式使两个相邻节点同色的可能性降到最小。不同的色彩代表了不同的频率，相邻节点色彩的不同，意味着能有效地降低频率的干扰。色彩波形优化彩色的最小数量，因此当阅读器保持一个配置成功的传输速率时，频率需要上升到预先定义的最大值。色彩波形是一个基于时分多址（TDMA）的方法，在这个方法中，每个阅读器根据分配的频率来随机选择一个时隙来传送。当碰撞时，阅读器选择一个新的时隙，并告知相邻节点它选择的频率。如果任何相邻节点调度使用了相同的频率（即相同的色彩）时，它将会对自己重新初始化，并选择一个新的频率，并告知相邻节点它的工作。选择频率的过程和调度时隙是可以根据需要重复的。如果成功传输的速率下降到先前设置的值，最大允许的色彩种类将会上升。色彩波形在一定网络规模内，减少了阅读器的碰撞。而在实际应用中，可以使用频率数是有限的。

分等级的 Q 学习（Hierarchical Q-learning，HiQ）算法[52]是一个分等级的在线学习算法，它可以找到阅读器碰撞问题的动态解决方案。HiQ 与色彩波形相似，通过随时间给每个阅读器分配频率来最小化阅读器的碰撞。但是，频率分配是通过学习相邻节点间的碰撞模式，然后分配最佳的频率。最佳的频率或是接近最佳的频率是通过重复环境的交互来实现的，这种环境交互即为碰撞模式和估计。HiQ 实现了三种基本的等级，有几种形式已经应用于实际的 RFID 系统，称之为阅读器（readers）、R-服务器（R-server）和 Q 服务器（Q-server）。最低的一级是 RFID 阅读器。RFID 阅读器在保留的时隙内使用预先调度的频率通信。阅读器检测相邻阅读器间（即有识别区域重叠的阅读器）的碰撞。这种先前已经解释过的统计性信息通过优化频率分配来实现。这种决定是上一级，即阅读器级服务器（R-servers）决定的。在低一级阅读器与 R-服务器之间存在着一对一的关系。Q 学习服务器（Q-server）在最高一级，并根据资源分配进行响应，通过基于表层的阅读器学习经验或是碰撞信息来寻找一个优化的调度方案。HiQ 学习算法分布于较低一级的阅读器中，并在最高一级进行集中化调度。HiQ 提供了优化的调度，导致了较高的吞吐量，但这是以集中化方案为代价的。

另外一些基于学习的方法包括神经网络和其他的遗传（Genetic）算法。神经网络方法是一个固定信道的解决方案，在这个解决方案中，每个神经代表着一个阅读器。系统限制和优化的标准被表示成能量函数，而神经网络试图最小化这些能量函数。这些神经的产物决定了信道对于代表的 RFID 阅读器是否是可用的。通用算

法是一种没有本地搜索的形式的算法。通用算法在分配信道时，一个简单的解决方案是为所有的 RFID 阅读器提供分配计划。通用算法使用一个解决方案集作为一个群，并开始演进处理。这个过程将重复执行，直到找到一个没有干扰的解决方案。

学习和调度方法是为了解决阅读器与阅读器间的碰撞。例如灯塔信道[53]、中心协调器[54]这样的交替过程用来解决多个阅读器与标签间的碰撞。脉冲分布协议[53]是基于灯塔机制的。当阅读器在它的识别区域读标签时，会使用控制信道周期性地广播一个灯塔信号。想要读标签的其他阅读器（即标签位于重叠的识别区内），会首先确认在它初始化自己的识别过程时，控制信道没有灯塔信号发送。后来的阅读器会等待，直到第一个阅读器执行完它的识别过程为止，因为在转换为自己的灯塔信号之前，控制信道需要保持空闲。不过，脉冲协议能够更好地消除由移动 RFID 阅读器而产生的碰撞。在中心协调（Central Cooperative，CC）方法[54]中，需要一个集中化的设备，以使标签数据复用到多阅读器。使用一个控制协调器，由于多个阅读器试图访问多个标签而产生的“多点对多点”的碰撞问题，可以转化成为“多点对一点”这样的典型碰撞问题。中心协调器把多个阅读器请求复用成为单一的一个标签请求。标签的响应分别对每一个阅读器进行解复用。中心协调器提供了一个处理阅读器与标签间碰撞的有效方法。但是，它需要一个具有传统阅读器所有功能的设备。

阅读器碰撞避免方案包括频率分配调度和在线学习方法。表 1-3 列出了解决碰撞问题的不同算法间的比较。不同的算法有其根本上的区别。例如，有些是中心控制的（即对信道分配和复用阅读器查询使用中心授权）。阅读器碰撞，特别是阅读器与阅读器之间的干扰，与传统无线网络中的频率分配问题是非常相似的。最根本的不同是，RFID 标签，特别是无源的 RFID 标签不能区别不同的阅读器。因此，两个阅读器与标签通信，必须使用不同的时隙，或是使用会话。根据 EPC Gen2 标准，标签需要为四个会话提供支持，允许两个以上的阅读器独立地识别标签。阅读器在分别的会话中，一步步选择性的独占标签，并完成把它们移动到其他会话中的动作。这样可以允许多个阅读器识别一般数量的标签。尽管这样有效，但是这需要标签额外的存储能力，从而增加了标签的成本。

表 1-3 阅读器碰撞算法比较

方法	算法	集中式控制	分布式控制	固定信道	动态信道
DCS[50]	调度	—	√	√	—
VDCS[51]	调度	—	√	—	√
HiQ[52]	学习	√	√	—	√
神经网络	学习	—	√	√	—
退火算法[56]	学习	√	—	√	—
遗传算法	学习	√	—	√	—

（续）

方法	算法	集中式控制	分布式控制	固定信道	动态信道
去冗余算法[55]	覆盖	—	√	—	—
CC[54]	覆盖	√	—	√	—

1.5　前景展望

本章的目的是对 RFID 的 MAC 协议提供一个最新的综述。详细叙述了解决阅读器和标签碰撞的不同方法。相关研究领域都是因为 RFID 系统作为基于射频的技术，使得 RFID 系统具有了独一无二的特性。这些特性包括特殊的交通模式（traffic pattern），以及在标签上能量和功能的限制。

在 RFID 系统的改进中，无线传感器网络（WSN）将扮演一个重要的角色。无线传感器网络在许多方式上，与传统的无线音频或是数据网络不同，但与 RFID 系统有许多的共同点。例如，在无线传感器中有大量的传感器节点，与有源 RFID 卡相似，这些节点是通过电池供电的。有时候传感器节点也会与无源卡一样，产生能量饥饿现象。另外，许多实际应用会大量部署传感器节点，所以无线传感器网络的密度会随时间和地理位置变化。这种特性与 RFID 系统部署托盘级（pallet-level）和物品级（item-level）的标签非常相似。更重要的是，与 RFID 系统一样，无线传感器网络是触发驱动的网络，并伴随着高级别的动态性。

但是，RFID 系统和无线传感器网络也存在着不同。传感器网络中，传感器节点常部署成 Ad Hoc 的形式，而不是部署成预先精心规划的形式，这通常使得节点可以自组织来监视和保持不同的网络功能性。而 RFID 系统中，并没有这种情况。RFID 标签，特别是无源 RFID 标签不能以 Ad Hoc 方式进行通信。基于这样的原因，无线传感器网络一般基本的 MAC 协议是 CSMA，而在 RFID 系统中，这种协议是不可行的。

参考文献［57-59］报道了关于无线传感器网络与 RFID 系统的集成。可能性的方案包括在标签上集成传感器，标签上集成无线传感器节点，在阅读器上集成无线传感器节点和无线设备，或是以 RFID 与传感器混合形式集成[59]。但是，这种集成对干扰观点提出了新的挑战。对于在 RFID 网络和无线传感器网络它们各自的领域中减少干扰，已经有了相当多的成果。尽管如此，随着无线传感器网络与 RFID 的集成，无线传感器节点、RFID 标签、RFID 阅读器这些设备的增加，情况将会变差，亟须有协作性的方案来解决设备之间的干扰。在不同的设备类型之间进行分布式的频率调度，RFID 标签用在帮助 MAC 协议的唤醒过程，节能方式下运行和无线传感器网络中上下文意识（context-aware）的 MAC 调度都是可能的方向。物品级（item-level）的标签，移动标签和移动传感器，以及在 RFID 和传感器网络的新领

域，上述这些都要求更复杂的 MAC 协议。

参考文献

1. D. Shih, P.L. Sun, D.C. Yen, and S.M. Huang, Taxonomy and survey of RFID anti-collision protocols, *Computer and Communications*, 29, 2150–2166, 2006.
2. Z. Tang and Y. He, Research of multi-access and anti-collision protocols in RFID systems, *2007 IEEE International Workshop on Anti-Counterfeiting, Security, Identification*, pp. 377–380, April 16–18, Xiamen, Fujian, China, 2007.
3. C. Mutti and C. Floerkemeier, CDMA-based RFID systems in dense scenarios: Concepts and challenges, *2008 IEEE International Conference on RFID*, pp. 215–222, April 16–17, Los Vegas, NV, 2008.
4. P. Vandenameele, *Space Division Multiple Access for Wireless Local Area Networks*, Kluwer Academic Publishers, Norwell, MA, 2001.
5. K. Finkenzeller, *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, John Wiley & Sons, Inc., England, U.K., 2003.
6. K. Ali, H. Hassanein, and A.M. Taha, RFID anti-collision protocol for dense passive tag environments, *LCN '07: Proceedings of the 32nd IEEE Conference on Local Computer Networks*, pp. 819–824, Dublin, Ireland, 2007.
7. J. Kim, W. Lee, E. Kim, D. Kim, and K. Suh, Optimized transmission power control of interrogators for collision arbitration in UHF RFID systems, *IEEE Communications Letters*, 11, 22–24, 2007.
8. W. Alsalih, K. Ali, and H. Hassanein, Optimal distance-based clustering for tag anti-collision in RFID systems, *33rd IEEE Conference on Local Computer Networks*, pp. 266–273, Montreal, QC, Canada, 2008.
9. J. Lee, T. Kwon, Y. Choi, S.K. Das, and K. Kim, Analysis of RFID anti-collision algorithms using smart antennas, *SenSys '04: Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 265–266, Baltimore, MD, 2004.
10. H. Vogt, Efficient object identification with passive RFID tags, *Pervasive '02: Proceedings of the First International Conference on Pervasive Computing*, pp. 98–113, Zurich, Switzerland, 2002.
11. B. Feng, J. Li, J. Guo, and Z. Ding, ID-binary tree stack anticollision algorithm for RFID, *ISCC '06: Proceedings. 11th IEEE Symposium on Computers and Communications*, pp. 207–212, Pula-Cagliari, Sardinia, Italy, 2006.
12. L. Bolotnyy and G, Robins, Randomized pseudo-random function tree walking algorithm for secure radio-frequency identification, *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 43–48, Buffalo, NY, 2005.
13. J. Capetanakis, Tree algorithms for packet broadcast channels, *IEEE Transactions on Information Theory*, 25, 505–515, 1979.
14. J.H. Choi, D. Lee, H. Jeon, J. Cha, and H. Lee, Enhanced binary search with time-divided responses for efficient RFID tag anti-collision, *ICC '07: IEEE International Conference on Communications*, pp. 3853–3858, Glasgow, Scotland, 2007.
15. F. Zhou, D. Jin, C. Huang, and M. Hao, Optimize the power consumption of passive electronic tags for anti-collision schemes, *Proceedings of 5th International Conference on ASIC*, vol. 2, pp. 1213–1217, Beijing, China, 2003.

16. C. Floerkemeier, Transmission control scheme for fast RFID object identification, *PerCom Workshops 2006: Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 457–462, Pisa, Italy, 2006.
17. L. Liu, Z. Xie, J. Xi, and S. Lai, An improved anti-collision algorithm in RFID system, *2nd International Conference on Mobile Technology, Applications and Systems*, pp. 137–142, Guangzhou, Guangdong, China, 2005.
18. N. Zhang and B. Vojcic, Binary search algorithms with interference cancellation RFID systems, *MILCOM 2005: IEEE Military Communications Conference*, vol. 2, pp. 950–955, Atlantic City, NJ, 2005.
19. M. Nanjundaiah and V. Chaudhary, Improvement to the anticollision protocol specification for 900 MHz class 0 radio frequency identification tag, *AINA '05: Proceedings of the 19th International Conference on Advanced Information Networking and Applications*, pp. 616–620, Taipei, Taiwan, 2005.
20. T. Hwang, B. Lee, Y.S. Kim, D.Y. Suh, and J.S. Kim, Improved anti-collision scheme for high speed identification in RFID system, *ICICIC '06: First International Conference on Innovative Computing, Information and Control*, vol. 2, pp. 449–452, Beijing, China, 2006.
21. T. Wang, Enhanced binary search with cut-through operation for anti-collision in RFID systems, *IEEE Communications Letters*, 10, 236–238, 2006.
22. S. Weis, S.E. Sarma, R.L. Rivest, and D.W. Engels, Security and privacy aspects of low-cost radio frequency identification systems, *Security in Pervasive Computing*, 2802, 201–212, 2004.
23. G. Khandelwal, A. Yener, and M. Chen, OPT: Optimal protocol tree for efficient tag identification in dense RFID systems, *ICC'06: IEEE International Conference on Communications*, vol. 1, pp. 128–133, Istanbul, Turkey, 2006.
24. S.H. Kim and P. Park, An efficient tree-based tag anti-collision protocol for RFID systems, *IEEE Communications Letters*, 11, 449–451, 2007.
25. W. Chen, S. Horng, and P. Fan, An enhanced anti-collision algorithm in RFID based on counter and stack, *ICSNC '07: Proceedings of the Second International Conference on Systems and Networks Communications*, pp. 21–24, Cap Esteral, French Riviera, France, 2007.
26. J. Myung, W. Lee, and J. Srivastava, Adaptive binary splitting for efficient RFID tag anti-collision, *IEEE Communications Letters*, 10, 144–146, 2006.
27. J. Myung and W. Lee, Adaptive binary splitting: A RFID tag collision arbitration protocol for tag identification, *Mobile Network Applications*, 11, 711–722, 2006.
28. J. Myung, W. Lee, J. Srivastava, and T.K. Shih, Tag-splitting: Adaptive collision arbitration protocols for RFID tag identification, *IEEE Transaction on Parallel and Distributed System*, 18, 763–775, 2007.
29. J. Ryu, H. Lee, Y. Seok, T. Kwon, and Y. Choi, A hybrid query tree protocol for tag collision arbitration in RFID systems, *ICC '07: IEEE International Conference on Communications*, pp. 5981–5986, Glasgow, Scotland, 2007.
30. F. Zhou, C. Chen, D. Jin, C. Huang, and H. Min, Evaluating and optimizing power consumption of anti-collision protocols for applications in RFID systems, *ISLPED'04: Proceedings of the 2004 International Symposium on Low Power Electronics and Design*, pp. 357–362, Newport, CA, 2004.

31. V. Namboodiri and L. Gao, Energy-aware tag anti-collision protocols for RFID systems, *PerCom '07: Fifth Annual IEEE International Conference on Pervasive Computing and Communications*, pp. 23–36, White Plains, NY, 2007.
32. N. Pastos and R. Viswanathan, A modified grouped-tag TDMA access protocol for radio frequency identification networks, *WCNC 2000: IEEE Wireless Communications and Networking Conference*, vol. 2, pp. 512–516, Chicago, IL, 2000.
33. C. Law, K. Lee, and K. Siu, Efficient memoryless protocol for tag identification, *DIALM '00: Proceedings of the 4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pp. 75–84, Boston, MA, 2000.
34. J. Choi, D. Lee, and H. Lee, Query tree-based reservation for efficient RFID tag anti-collision, *IEEE Communications Letters*, 11, 85–87, 2007.
35. K.W. Chiang, C. Hua, and P. Yum, Prefix-length adaptation for PRQT protocol in RFID systems, *GLOBECOM'06: IEEE Global Telecommunications Conference*, pp. 1–5, San Francisco, CA, 2006.
36. K.W. Chiang, C. Hua, and T.P. Yum, Prefix-randomized query-tree protocol for RFID systems, *ICC '06: IEEE International Conference on Communications*, vol. 4, pp. 1653–1657, Istanbul, Turkey, 2006.
37. J. Cha and J. Kim, Dynamic framed slotted ALOHA algorithms using fast tag estimation method for RFID system, *CCNC 2006: 3rd IEEE Consumer Communications and Networking Conference*, vol. 2, pp. 768–772, Los Vegas, NV, 2006.
38. J. Cha and J. Kim, Novel anti-collision algorithms for fast object identification in RFID system, *Proceedings 11th International Conference on Parallel and Distributed Systems*, vol. 2, pp. 63–67, Cambridge, MA, 2005.
39. J. Park, M.Y. Chung, and T. Lee, Identification of RFID tags in framed-slotted ALOHA with robust estimation and binary selection, *IEEE Communications Letters*, 11, 452–454, 2007.
40. J. Park, M.Y. Chung, and T. Lee, Identification of RFID tags in framed-slotted ALOHA with tag estimation and binary splitting, *ICCE '06: First International Conference on Communications and Electronics*, pp. 368–372, San Diego, CA, 2006.
41. G. Khandelwal, A. Yener, K. Lee, and S. Serbetli, ASAP: A MAC protocol for dense and time constrained RFID systems, *ICC '06: IEEE International Conference on Communications*, vol. 9, pp. 4028–4033, Istanbul, Turkey, 2006.
42. S. Lee, S. Joo, and C. Lee, An enhanced dynamic framed slotted ALOHA algorithm for RFID tag identification, *MobiQuitous 2005: The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pp. 166–172, San Diego, CA, 2005.
43. W.J. Shin and J.G. Kim, Partitioning of tags for near-optimum RFID anti-collision performance, *WCNC 2007: IEEE Wireless Communications and Networking Conference*, pp. 1673–1678, Hong Kong, China, 2007.
44. C. Floerkemeier, Bayesian transmission strategy for framed ALOHA based RFID protocols, *IEEE International Conference on RFID 2007*, pp. 228–235, Grapevine, TX, 2007.
45. J. Choi, D. Lee, and H. Lee, Bi-slotted tree based anti-collision protocols for fast tag identification in RFID systems, *IEEE Communications Letters*, 10, 861–863, 2006.

46. J. Kim, J. Yu, J. Myung, and E. Kim, Effect of localized optimal clustering for reader anti-collision in RFID networks: Fairness aspects to the readers, *ICCCN 2005: 14th International Conference on Computer Communications and Networks*, pp. 497–502, San Diego, CA, 2005.
47. C.P. Wong, Grouping based bit-slot ALOHA protocol for tag anti-collision in RFID systems, *IEEE Communications Letters*, 11, 946–948, 2007.
48. D.W. Engels and S.E. Sarma, The reader collision problem, *2002 IEEE International Conference on Systems, Man and Cybernetics*, vol. 3, pp. 641–646, Hammamet, Tunisia, 2002.
49. K.S. Leong, M.L. Ng, and P.H. Cole, The reader collision problem in RFID systems, *MAPE 2005: IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, vol. 1, pp. 658–661, Beijing, China, 2005.
50. J. Waldrop, D.W. Engles, and S.E. Sarma, Colorwave: An anticollision algorithm for the reader collision problem, *ICC '03: IEEE International Conference on Communications*, vol. 2, pp. 1206–1210, Anchorage, AK, 2003.
51. J. Waldrop, D.W. Engles, and S.E. Sarma, Colorwave: A MAC for RFID reader networks, *WCNC 2003: 2003 IEEE Wireless Communications and Networking*, vol. 3, pp. 1701–1704, New Orleans, LA, 2003.
52. J. Ho, D.W. Engels, and S.E. Sarma, HiQ: A hierarchical Q-learning algorithm to solve the reader collision problem, *SAINT Workshops 2006: International Symposium on Applications and the Internet Workshops*, pp. 88–91, Phoenix, AZ, 2006.
53. S.M. Birari and S. Iyer, Mitigating the reader collision problem in RFID networks with mobile readers, *13th IEEE International Conference on Networks*, pp. 463–468, Kuala Lumpur, Malaysia, 2005.
54. D. Wang, J. Wang, and Y. Zhao, A novel solution to the reader collision problem in RFID system, *WiCOM 2006: International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, Wuhan, Hubai, China, 2006.
55. B. Carbunar, M.K. Ramanathan, M. Koyuturk, C. Hoffmann, and A. Grama, Redundant reader elimination in RFID systems, *IEEE SECON 2005: 2005 Second Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, pp. 176–184, Santa Clara, CA, 2005.
56. C. Lin and F. Lin, A simulated annealing algorithm for RFID reader networks, *WCNC'2007: IEEE Wireless Communications and Networking Conference*, pp. 1669–1672, Hong Kong, China, 2007.
57. J. Cho, Y. Shim, T. Kwon, Y. Choi, and S. Pack, SARIF: A novel framework for integrating wireless sensor and RFID networks, *IEEE Wireless Communications*, 14, 50–56, 2007.
58. L. Zhang and Z. Wang, Integration of RFID into wireless sensor networks: Architectures, opportunities and challenging problems, *Fifth International Conference on Grid and Cooperative Computing Workshops*, pp. 463–469, Changsha, Huan, China, 2006.
59. H. Liu, M. Bolic, A. Nayak, and I. Stojmenovic, Integration of RFID and wireless sensor networks, *Proceedings of The First ACM Workshop on Convergence of RFID and Wireless Sensor Networks and Their Applications*, Sydney, Australia, 2007.

第2章 RFID的防碰撞算法

在无线射频识别（RFID）系统中，标签存储着全球唯一的识别号，并附着在被识别物品上。阅读器通过给附有识别标签的物品发射无线射频信号来识别物品。像其他无线通信系统一样，RFID系统也存在信号干扰问题。主要有两类信号干扰，一类被称为阅读器碰撞，它存在于多个阅读器同时发射信号来识别同一个标签的时候；另一类称为标签碰撞，它存在于多个标签同时响应一个阅读器的时候。碰撞隐藏和减慢了标签的识别过程。因此，需要标签防碰撞协议和阅读器防碰撞协议来分别减少标签碰撞和阅读器的碰撞，以便于提高识别过程的性能。本章将主要介绍一些已存在的阅读器防碰撞和标签防碰撞协议。我们不仅对协议进行了广泛的调查，还对协议提出了新的研究方向。

2.1 概述

RFID系统的前端由标签和阅读器两部分组成[1]。在RFID系统中，标签存储着全球唯一的识别号，并附着在被识别物品上。阅读器发射射频信号通过识别标签间接地识别标签所附着的物品。标签被设计用于全球范围内的商业及其他相关应用领域的部署，标签应设计得小巧，低成本，并集成有限计算和通信能力的简单电路[2]。大部分RFID标签是被动的，标签上没有电源，只能接收从阅读器产生的射频信号来产生能量。当标签和阅读器足够靠近时，它们之间可以互相通信。在这种情况下，可以说标签在阅读器的识别区域内。像其他无线通信系统一样，RFID系统也存在着信号干扰问题[3]。信号干扰主要有两类，一类被称为阅读器碰撞，它存在于多个阅读器同时发射信号来识别标签的时候；另一类称为标签碰撞，它存在于多个标签同时响应一个阅读器的时候。碰撞隐藏和减慢了标签的识别过程。因此，需要阅读器防碰撞协议和标签防碰撞协议分别来减少阅读器碰撞和标签碰撞，以提高识别过程的性能。本章将主要介绍一些已存在的阅读器防碰撞和标签防碰撞协议。我们不仅对协议进行了广泛的调查，还对协议提出了新的研究方向。

由于标签是从阅读器得到能量的，标签的响应范围（也被称为识别范围）比阅读器的射频信号的传输范围（也被称为干涉范围）要小得多[4]。此外，标签和阅读器具有不同的计算和通信能力。鉴于这样的不对称性，我们不能依靠通常在无线局域网[3]中使用的RTS/CTS的碰撞避免机制来解决碰撞问题。

一些阅读器的防碰撞议是基于时分多址（TDMA）、频分多址（FDMA）或是载波监听多路访问（CSMA）技术来减少阅读器的碰撞。基于TDMA的阅读器防碰

撞协议把传输的时间分成多个间隔，一个阅读器只能在它所分配的间隔内传输信息，可以采用分布式或是集中式的方式来分配时间间隔[6]。Waldrop 等人提出了两种分别为分布式色彩选择（DCS）和 Colorwave 的分布式的基于 TDMA 的阅读器防碰撞协议。首先引入一个阅读器图，在图中任何的两个阅读器被定为相邻的，并且互相之间有一个边界，在这个边界中，阅读器可能互相会产生干扰。每个阅读器被分配一个颜色，此颜色表示阅读器发送信号的具体时隙。如果所有的相邻节点都有不同的颜色，阅读器碰撞将会避免。在 DCS 协议中，最大的色彩数量是固定的，阅读器只在它分配的颜色（即时隙）内发送数据。相反的，Colorwave 协议具有动态的最大颜色数量值，并且具有动态的颜色分配机制来最小化阅读器图中颜色需要的数量。随着使用的颜色数量的减少，信息传送的效率将会增加。

基于 FDMA 的协议把所有可用的频率带划分成多个频率信道。如果一个频率信道一次只分配给一个收发器，多个收发器将会无干扰地同时收发数据。Ho 等人提出了同时基于 TDMA 和 FDMA 的 HiQ 协议[8]。HiQ 协议试图了解阅读器的碰撞模式，以及在时间上有效分配频率来最小化阅读器碰撞的产生。HiQ 协议使用称为 Q-learning 的分布式的、分等级的，并且是在线学习的方案来决定频率和时间的分配。通过重复地与系统进行交互，Q- learning 试图在时间上寻找一种优化的频率分配方案。EPCglobal Gen2[9]是一种著名的采用 FDMA 技术来解决阅读器碰撞的协议。阅读器可以通过跳频、扩频技术来选择一个独立的收发信道避免干扰。

CSMA 是另外一个用来解决阅读器碰撞的机制。在 CSMA 机制中，每个阅读器在发送信息之前需要检查载波（被分享的通信信道）是否空闲。如果载波被检测到是空闲的，阅读器会立即发送信息。否则，阅读器会延迟一个随机的时间周期，然后再次检测载波。欧洲通信标准组织（ETSI）EN 302 208 标准[10]使用基于 CS-MA 的“先听后说”（Listen Before Talk，LBT）的机制来解决阅读器碰撞。

多个标签防碰撞协议已经被提出，用来减少标签的碰撞。主要可以分为三类：基于 ALOHA 的、基于树的和基于计数的协议[11]。ALOHA[12]、时隙 ALOHA[13]和帧时隙 ALOHA[14]协议是基于 ALOHA 的协议。在 ALOHA 协议中，阅读器首先发送一个命令促使标签发送回它们的 ID 号。接收到阅读器的识别信号后，在识别区中的每个标签会等待一个随机的回应时间，把它们自己的 ID 号传回给阅读器。如果在一个标签的响应期间，没有碰撞发生，这个标签的 ID 号就能够被正常识别。在时隙 ALOHA 协议中，随机的回应时间必须是一个预先设置的时隙时间。帧时隙 ALOHA 协议与时隙 ALOHA 协议相似，区别在于，帧时隙 ALOHA 协议的识别过程被分为一系列的帧，每个帧具有一个固定的时隙，标签只能在一个帧周期内选择一个随机的时隙来发送它的 ID 号。基于 ALOHA 的协议较简单，但是存在着标签饥饿问题，即当某个标签的响应总是与其他标签的响应碰撞时，它的 ID 号就可能永远不会被正确地识别。

基于树的协议[15-19]的基本思想是根据标签的 ID 号将遇到碰撞的标签分成几个

子群，直到每个子群中只有一个标签能成功地被识别为止。这些协议既可以应用到有可写存储单元的标签中，也可以应用到无可写存储单元的标签中。有存储单元的标签需要更高的成本。但是，协议应用在具有存储能力的标签中，可以有更好的性能。查询树（Query Tree，QT）协议[16]应用在没有可写存储单元的标签中。在此协议中，阅读器广播一个具有可变长的字符串S请求位给标签。如果标签的ID前缀与S匹配，此标签则会把它的ID号发送给阅读器。当碰撞发生时，阅读器会广播一个更长的字符串S0或者S1，把碰撞的标签分裂成两个子群。位与位的二进制树的协议[15]应用于具有可写的存储单元的标签。在此协议中，阅读器首先广播一个请求命令，每个标签会响应它的ID号的第一位。如果碰撞发生，阅读器将会用0（或1）来识别标签。只有那些第一位是0（或1）的标签才会把它接下来的位发送给阅读器。在这种方式下，标签将会持续地分裂成两个组。其他基于树的协议，例如EPCglobal Class0[19]、树隙ALOHA（Tree-Slotted ALOHA，TSA）[17]、二进制时隙查询树算法（Bi Slotted Query Tree Algorithm，BSQTA）和二进制时隙碰撞跟踪树算法（Bi Slotted Collision Tracking Tree Algorithm，BSCTTA）[18]，也具有相似的原理来分裂标签，解决标签的碰撞。基于树的协议的主要缺点是它的性能受标签ID号的长度和分布的影响。通常情况下，基于树的协议需要比基于ALOHA协议[20]更长的识别时间，但是基于树的协议不存在标签饥饿问题。

基于计数器的协议[11,20-23]与基于树的协议相似。主要的不同在于，前者依靠静态的ID号来分裂，后者则是依靠动态改变的计数器来进行分裂标签。ISO/IEC 18000-6B[22]是一种标准的采用基于计数器的标签防碰撞协议。在ISO/IEC 18000-6B中，每个标签有一个初始化设置为0的计数器。当阅读器发送请求信号给标签时，每个计数器值为0的标签会发送它的ID号给阅读器。当碰撞发生时，计数器值大于0的标签会把它的计数器值再加1，计数器值为0的标签会随机地产生一个随机位0或1，并把此0或1加到它们的计数器当中。采用这种方式，计数器值为0的标签会分裂成两个子群。其他基于计数器的协议，例如自适应性的二进制分裂（Adaptive Binary Splitting ABS）协议[24]，使用相似的原理来分裂标签遇到的碰撞。基于计数器的协议不存在标签饥饿问题。此外，基于计数器的协议具有稳定的特性，因为它们的性能不会受标签ID号的长度或是标签ID号的分布的影响。

本章中所提及的防碰撞协议和它们的分类见表2-1。阅读器防碰撞协议分为TDMA、FDMA和CSMA三类。标签防碰撞协议分为基于ALOHA、基于树和基于计数器三类。所有表2-1中提及的协议将会在接下来的各节中进行详细的介绍。

本章接下来各部分中，2.2节首先定义碰撞问题，2.3节将会详细介绍TDMA、FDMA和CSMA三类阅读器防碰撞协议，2.4节会详细介绍包括基于ALOHA、基于树和基于计数器在内的标签防碰撞协议，同时在2.3节和2.4节中，将会给出一些例子来更好地理解上述协议。最后，在2.5节中给出总结和防碰撞新的研究方向的一些建议。

表 2-1　本章将要讨论的协议

碰撞类型	分类	协　议
阅读器碰撞	TDMA	DCS[7]
		Colorwave[7]
	FDMA	HiQ[8]
		EPCGlobal Gen 2[9]
	CSMA	ETSI 302 208 Standard[10]
标签碰撞	基于 ALOHA	ALOHA[12]
		时隙 ALOHA[13]
		帧时隙 ALOHA[14]
		ISO/IEC 18000-6A[22]
	基于树	QT[16]
		Bit-by-bit binary tree（渐进二叉树）[15]
		EPCGlobal Class 0[19]
		TSA[17]
		BSQTA[18]
		BSCTTA[18]
		AQS（Adaptive Query Splitting，自适应查询分裂）[30]
	基于计数器	ISO/IEC 18000-6B[22]
		ABS[30]

2.2　RFID 系统的阅读器碰撞问题

当一个阅读器（也称为识别器）发送一个请求信号给标签时，它也会提供能量给无源标签。如果阅读器和无源标签距离足够近，阅读器可以接收到标签发送来的信号。在此情况下，可以说，标签在阅读器的识别区域内。当两个以上的阅读器距离非常近，或者是在一个阅读器的识别区域内，存在许多标签，干扰将会增加，这些干扰主要分为阅读器碰撞和标签碰撞。接下来，将会描述这两种碰撞。

1. 阅读器碰撞（或称为阅读器干扰）问题

由于标签是由阅读器提供能量的，标签的响应区（比如识别区）比阅读器的发送区（或称为干扰区）要小得多。当一个标签在阅读器 A 的识别区内，而在阅读器 B 的干扰区内时，由于阅读器间的干扰，标签不能正确地接收来自阅读器 A 的请求命令，或者是阅读器 A 不能正确地接收（解读）来自标签的响应，这被称为阅读器碰撞。如图 2-1 所示，标签 T 在阅读器 A 的识别区内，而在阅读器 B 的干扰区内，在这种情况下，阅读器碰撞将有可能会发生。

2. 标签碰撞问题

为了识别在识别区内的标签，阅读器会发出一个请求信号要求标签发回它的ID号。当阅读器的识别区内多个标签同时响应阅读器的请求时，碰撞将会发生，从而阅读器不能正常地识别标签，这被称为标签碰撞。如图2-1所示，标签S和标签T在阅读器A的识别区内。如果标签S和标签T响应阅读器A的请求后同时发射它的ID号，标签碰撞将会发生，阅读器A不能识别到标签S和标签T中的任何一个标签。

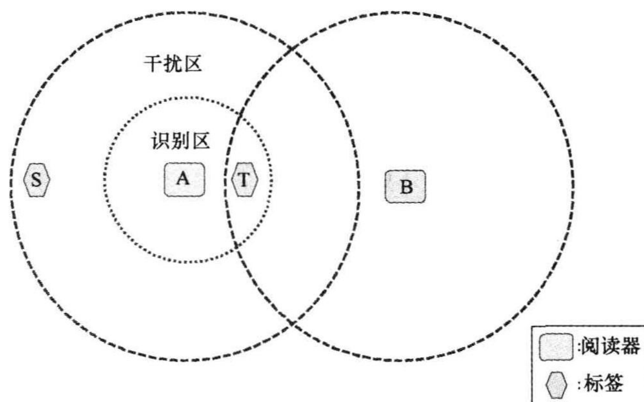


图2-1 识别区与干扰区之间的关系

2.3 阅读器防碰撞协议

数个阅读器防碰撞协议已经被提出来解决阅读器的碰撞问题。它们分为三类：TDMA、FDMA和CSMA协议[5]。接下来，将详细地介绍阅读器防碰撞协议。

2.3.1 TDMA协议

基于TDMA的阅读器防碰撞协议的基本思想是把整个时间周期分成若干个时间间隔，并允许一个阅读器只能在它所分配的时间间隔内发送信息。在这种方式下，阅读器碰撞将可以避免。接下来将介绍两个基于TDMA的阅读器防碰撞协议：DCS和Colorwave算法[7]。

2.3.1.1 DCS算法

DCS是Waldrop在参考文献［7］中提出的阅读器防碰撞协议。时隙假设是被周期性地标记为0、1和最大色彩的颜色。DCS通过提出一个阅读器图来解决阅读器的碰撞。在阅读器图中，阅读器被表示成节点或是相邻的节点，并且存在一个节点间是否存在干扰的边界。随后，使每个阅读器标记上一个色彩，来表示传输信号的一个具体时隙。如果所有相邻的节点具有不同的色彩，则阅读器的碰撞就能避免。

DCS是一个分布式的算法，可以从阅读器相互之间设置的0、1和最大色彩随机地、局部性地选择一个色彩（即时隙），其中最大色彩是一个输入参数，它的值将不会改变。当一个阅读器想要发送信息给标签时，它会把信息放入队列中，直到此阅读器选择的色彩可以使用为止。如果阅读器在它选择的时隙内发送信息时，发现存在着碰撞，阅读器则会重新选择一个新的色彩，并通知相邻阅读器相应地改变它们的色彩。需要注意的是，DCS算法需要同步时隙的时间，但是不必与系统内所有阅读器的色彩值同步。

2.3.1.2 Colorwave算法

Colorwave算法［或者说是可变最大值的分布式色彩选择（Variable-Maximum Distributed Color Selection，VDCS）算法］是一个DCS的扩展算法。在Colorwave算法中，提出了优化需要标记的阅读器图的色彩数量（比如最大色彩）的机制。如果使用的色彩减少，信号传输的效率将提高。

当阅读器自己观察到，或是被相邻阅读器检测到数据传输成功的概率小于最大色彩阈值时，它就会增加本地最大色彩值，同时给相邻阅读器广播这个新的最大色彩值，使相邻阅读器也重新选择色彩来减少传输的碰撞。相反地，当数据传输成功的概率大于最大色彩阈值时，阅读器会减少本地最大色彩值以减少传输等待时间。

2.3.2 FDMA协议

FDMA协议把所有可用的频率带划分成若干个互相不干扰的信道。阅读器可以使用不同的信道来同时与标签进行通信。接下来，将介绍两个协议：HiQ[8]和EPC-global Gen 2[9]协议，这两个协议都采用了FDMA机制来解决阅读器碰撞。

2.3.2.1 HiQ协议

HiQ协议[8]是一个基于TDMA和FDMA的、分级的、分布式的、在线学习的算法，来解决阅读器碰撞。设计的目标是最大化在阅读器和标签之间并发通信的信道数量，并通过学习阅读器的碰撞模式来最小化阅读器碰撞的数量，以使得有效地将每个时隙分配给阅读器。

如图2-2所示，HiQ分级的控制结构由阅读器、R-server和Q-server组成。RFID阅读器在最低的一级，在R-server级的每个服务器管理着多个阅读器。当某个阅读器需要发送信息给它的识别区域内的标签时，它必须首先从它的主R-server处请求资源，即频率信道和时隙。阅读器只有在主R-server分配于一个时隙内的具体频率信道之后，才能够发送信息。

在这样的分布式架构中，相邻阅读器可以在相同的时隙或者相同的频率信道内发送信息，从而会造成碰撞。阅读器需要检测与相邻节点之间的碰撞。每个阅读器需要报道碰撞的数量、碰撞的类型和成功读它的主R-server的次数。随后，R-server能够根据反馈的报告判断哪些从阅读器间有相互干扰，并重新动态地分配资源来避免碰撞。

R-server能够分配的资源来自分级结构中的主Q-server（Q-learning server）。由于较好的灵活性和可扩展性，Q-server在分级架构中可以较好地完成自己工作。但是，在整个系统中，只有一个根Q-server控制所有的频率信道和时隙的分配。

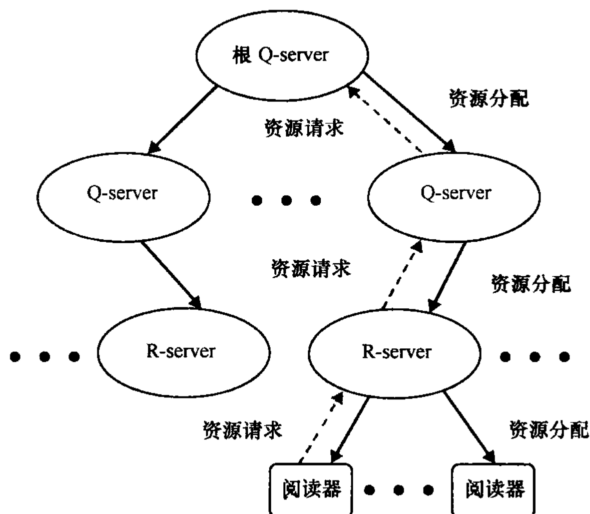


图2-2　HiQ协议的分级控制结构

2.3.2.2　EPCglobal Gen 2协议

Class 1 Generation 2 UHF标准是EPCglobal提出的，使用FDMA技术来减少阅读器的干扰。整个分配频带被分成若干个信道，一个阅读器只能使用一个信道来进行通信。阅读器和标签分开使用载波频率。即阅读器（或标签）将只会与阅读器（或标签）发生碰撞。阅读器使用跳频扩频技术来避免干扰。在欧洲，频率分配时规定为200kHz的带宽[25]。建议阅读器使用偶数的信道，而标签的散射信号使用奇数信道。在美国，频率分配时规定为500kHz的带宽。所有的信道对阅读器是有效的，而标签在这些信道的边界处散射。EPCglobal Gen 2协议能够解决阅读器的碰撞。因为大部分低成本的阅读器没有频率选择能力，标签碰撞的问题仍然存在[3]。

2.3.3　CSMA协议

CSMA协议在有线或是无线系统中，是一种用来避免碰撞的常规机制。在这个机制中，每个设备在发射信息前，都需要检查信道是否空闲。如果信道是忙的，那么设备会等待，直到信道空闲为止。

ETSI 302 208是欧洲采用称为“LBT”的CSMA机制来解决阅读器碰撞的标准。对于RFID的应用，它分配865～868MHz的频率，把这频带划分成15个信道，其中，每个信道为200kHz的带宽。当最大有效的辐射能量为2W时，只有其中10个信道是用来通信的，另外5个信道被定义为保护带，或是为低能量的阅读器保

留。在发射数据前，阅读器的接收模块首先被激活，在一个具体的时间间隔内（5ms）监视可选的信道。如果在此特定时间间隔内检测到信道都是空闲的，那么阅读器将会在 4s 内发送它的数据，然后阅读器会激活接收模块来检测信号的干扰。如果信道被其他阅读器使用，那么阅读器会选择其他空闲的信道来传输数据。

2.4 标签防碰撞协议

为了减少标签的碰撞，多个标签防碰撞协议已经被提出。这些协议主要可以分为 3 类：基于 ALOHA 的、基于树的和基于计数器的协议[11]。接下来，将详细地介绍这些协议。

2.4.1 基于 ALOHA 的协议

基于 ALOHA 的标签防碰撞协议[21,26-28]是根据运行于概率方式下的一种后退机制。基于 ALOHA 的协议试图错开识别区域内标签的响应时间。接下来，将介绍几个基于 ALOHA 的协议：ALOHA[12]、时隙 ALOHA[13]和帧时隙 ALOHA[14]协议。通常，基于 ALOHA 的协议比较简单，且性能比较正常。但是它们会出现标签饥饿问题，因此由于标签的响应会相互碰撞导致它可能永远无法被识别。

2.4.1.1 ALOHA 协议

ALOHA 协议[12]是最简单的基于 ALOHA 的标签防碰撞协议。当阅读器请求标签发回它的 ID 号时，在识别区内的每个标签会自己选择一个随机的回退时间，在这个回退时间之后，把标签的 ID 号发送给阅读器。如果在标签 ID 号的发送期间没有碰撞发生，此 ID 号将会被阅读器成功地识别，被识别 ID 号的标签将停止对阅读器的响应。否则，标签会重复地选择一个随机的后退时间，发送它的 ID 号，直到 ID 号被阅读器识别为止。

2.4.1.2 时隙 ALOHA 协议

在时隙 ALOHA 协议[13]中，随机的回退时间必须是多个预先设置的时隙时间。需要注意的是，时隙时间经常被设置成一个时间周期，这个时间周期必须足够长，可以使标签发送完它的 ID 号，并使阅读器识别这个 ID 号。阅读器需要为识别区内的所有标签同步时隙时间。如果在一个时隙时间内，只有一个标签传输它的 ID 号，那么阅读器可以正确地识别到。没有被阅读器识别到的标签将会再次选择一个随机的时隙来发送它的 ID 号。如参考文献［29］中报道的，时隙 ALOHA 协议的性能是 ALOHA 协议的两倍，因为部分标签 ID 响应的碰撞没有在时隙 ALOHA 协议中发生。

2.4.1.3 帧时隙 ALOHA 协议

帧时隙 ALOHA 协议[14]中，整个识别过程被分成一系列的帧，每个帧具有多个时隙。在接收阅读器的请求命令时，每个标签只在某个帧期间随机选择的一个时隙

内响应。如果在时隙中只有一个标签响应，阅读器可以成功地识别标签。没有成功被识别的标签将会在下一个帧中再次选择一个时隙来发送它的 ID 号。当没有标签响应时，即表示所有的标签都已经被成功地识别。帧的循环会持续进行，直到标签都被成功地识别。

如图 2-3 所示，是一个帧时隙 ALOHA 协议的例子，在这个例子中，每个帧有四个时隙。假设，在一个阅读器识别区内，有六个标签，每个标签具有唯一的五位 ID 号。协议的执行过程描述如下。

1）阅读器首先发送请求命令来同步帧的起始位。

2）接收请求命令后，标签会随机的选择帧 0 中的四个可用时隙中的一个来发送回它的 ID 号。在这个例子的帧 0 中，只有在时隙 1 中 ID 号为 01110 的标签能够被成功地识别。碰撞在时隙 2 和时隙 4 中发生，在时隙 3 中没有标签响应。

3）已经被识别的标签会使用选择命令来读或是写数据。在接下来的帧中，它会停止对请求命令的响应。

4）如图 2-3 中的帧 1 和帧 2，阅读器重复地发送请求命令，直到所有的标签被成功地识别为止。

帧时隙 ALOHA 协议的一个不足之处是，当时隙的数量与识别区内标签的数量不是很好地匹配时，它的性能将会降低。动态帧时隙 ALOHA 协议[26-28,30]根据估计的标签数量来动态调节帧的大小，来试图消除这个不足。它的性能优于帧时隙 ALOHA 协议的性能。

2.4.1.4 ISO/IEC 18000-6A 协议

ISO/IEC 18000-6A[22]标准定义为使用 860 ~ 960MHz 通信的 RFID 系统中的空中干扰。此标准中定义了分别为 A、B 和 C 的三种类型的通信协议。其中，A 和 C 类型是基于 ALOHA 的协议。由于 C 协议是 A 协议演变而来，所以此处只介绍 A 协议。

ISO/IEC 18000-6A 中，阅读器通过发送初始化循环命令来初始化识别过程的一个循环。在这个命令中，指出了代表一个循环大小的时隙数。需要注意的是，阅读器能够根据此次循环中碰撞的数量来动态地确定下一轮循环的适合的循环大小。接收到命令之后，标签会随机地选择一个时隙来发送它的 ID 号给阅读器。标签会使用一个时隙计数器来跟踪当前的时隙。当被选择的时隙到达时，标签会等待一个在 0 ~ 7 周期范围内的随机延时，并使用一个随机选择的四位标签签名来响应。如果只有一个标签响应，且这个标签的签名被阅读器正确地接收，阅读器则会发送包含签名的下一个时隙命令给标签进行确认。否则，将会发送关闭时隙命令。标签具有以下行为。

标签如果在当前时隙内没有响应，并且接收到命令是下一个时隙或是关闭时隙命令，则时隙计数器将会增加 1。

如果标签在当前隙内响应，而且接收到的命令是关闭时隙，则标签的时隙计数

	帧0					帧1					帧2				
		时隙1	时隙2	时隙3	时隙4		时隙1	时隙2	时隙3	时隙4		时隙1	时隙2	时隙3	时隙4
阅读器	请求					请求					请求				
标签1			10010						10010						
标签2		01110													
标签3					00101		00101							00101	
标签4			11011				11011						11011		
标签5			10110					10110							
标签6					01001					01001					
状态		成功	碰撞	空闲	碰撞		碰撞	成功	成功	成功		空闲	成功	成功	空闲

图2-3 帧时隙ALOHA协议的一个例子

器将会增加1。

如果标签在当前隙内响应，并且接收到的命令是同一个签名的下一个时隙，标签将会转为安静状态。

在一个循环期间，通过发送等待循环命令给标签，阅读器能够暂停循环。循环的等待允许阅读器处理已选择标签的读写数据的对话。当时隙的计数与初始化循环命令的循环大小相等时，循环将结束，所有不处于静止状态的标签（例如标签仍然没有被识别）将会随机的选择一个新的时隙和新的随机签名来进入一个新的循环。

2.4.2 基于树的协议

基于树的标签防碰撞协议的基本思想是根据标签的ID号，把遇到碰撞的标签重复分裂成数个子群，直到在一个子群中，只有一个标签能够被成功地识别为止。这些协议既可以被应用到具有可写的存储器的标签中，也可以应用到不具有可写的存储器的标签中。有存储器的标签需要更高的成本。不过，这样的标签也具有更高的性能。通常情况下，基于树的协议比基于ALOHA的协议需要更长的识别标签的时间，但是，基于树的协议不存在标签饥饿问题。基于树的协议的另一个不足是它的性能会受标签ID号的长度和分布的影响。接下来，将介绍几类基于树的协议：查询树[16]、二进制树[15]、EPCglobal Class 0[19]、TSA[17]、BSQTA[18]和BSCTTA[18]协议。

2.4.2.1 查询树协议

在查询树（QT）协议[16]中，阅读器首先广播一个位字符串S的请求给标签。ID号前缀与S匹配的标签将会把它的整个ID号发给阅读器。如果一次只有一个标签响应，此标签则可以被成功地识别。但是，如果有多个标签同时响应时，碰撞将会发生。在这种情况时，阅读器会再次广播一个在字符串S后多一位0或是1的更长的位字符串，即S0或是S1。显然，具有S前缀的标签将会被分裂成S0和S1两个子群。分裂的过程将会重复地发生，直到每个识别区内的标签都能被成功识别为止。查询树协议是一个跟存储无关的协议，因为它不需要标签具有额外的可写的存储器。可以发现，标签ID号的长度和分布会影响查询树协议的识别延迟。具体地说，如果标签具有连续的ID号，为了识别这些标签，请求位字符串会变得很长。识别过程的延迟时间也会随之变得很长。

下面，将展示一个查询树协议的例子。假设，存在六个标签，它们的ID号分别是0010、0011、1001、1100、1101和1110。查询树协议的标签识别过程描述如下。

1）阅读器首先发送一个S=“0”的请求位字符串，并把另一个请求位字符串“1”压入堆栈。ID号为0010和0011的标签的第一位与S匹配。它们会同时把它们的ID号发送给阅读器，从而会产生碰撞。

2）阅读器随后发送一个更长的字符串 S＝“00”，并把“01”压入堆栈。ID 号为 0010 和 0011 的标签会同时响应请求，碰撞会再次发生。

3）然后阅读器发送一个字符串 S＝“000”，并把“001”压入堆栈。没有一个标签的前缀与 S 匹配，所以没有响应发生。

4）由于没有响应，标签会从堆栈中弹出“001”，并它作为请求位字符串发送出去。ID 号为 0010 和 0011 的标签会同时响应这个请求，碰撞又会发生。

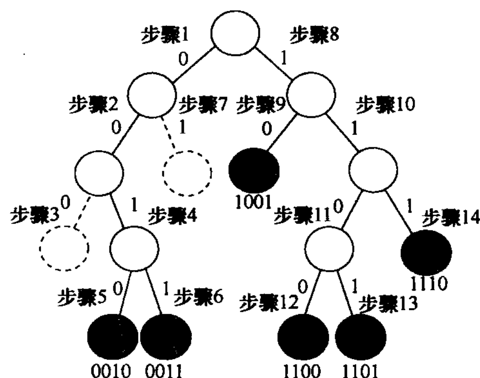
5）阅读器发送一个 S＝“0010”的位字符串，并把“0011”压入堆栈。只有 ID 号为 0010 的标签会响应此请求，然后此标签可以被顺利地识别。

6）在标签被成功识别之后，阅读器会从堆栈中弹出“0011”，并把它作为请求位字符串发送。只有 ID 号为 0011 的标签会响应请求，然后此标签被成功地识别。

识别的过程会被重复地执行，直到堆栈为空为止。最后，所有的标签都将会被成功地识别。表 2-2 为整个过程的步骤和相关的树图。

表 2-2　QT 协议的识别过程

步　骤	请求比特串 S	响　应
1	0	碰撞
2	00	碰撞
3	000	空
4	001	碰撞
5	0010	0010
6	0011	0011
7	01	空
8	1	碰撞
9	10	1001
10	11	碰撞
11	110	碰撞
12	1100	1100
13	1101	1101
14	111	1110



2.4.2.2　逐位二进制树协议

使用具有可写入能力的存储器，二进制树协议[15]能够有效地降低标签的碰撞。在这个协议中，阅读器首先广播一个请求命令，每个标签将会把它的 ID 号的第一位发送给阅读器。如果碰撞发生，阅读器会使用 0（或是 1）来确认标签。只有第一位是 0（或是 1）的标签才会把它的下一位发送给阅读器。上述过程一直重复，

直到只有一个标签响应为止。为了识别，阅读器会要求标签发送它 ID 号剩余的位。具有存储能力的标签，可以跟踪识别过程和响应的状态。与查询树协议不同，二进制树协议不需要阅读器发送长的 ID 号前缀，阅读器一次只发送一位。所以，识别过程的延迟将会减少。

2.4.2.3 EPCglobal Class 0

在 EPCglobal Class 0[19]中，标签会响应阅读器的请求，并把它 ID 号的第一位发送给阅读器。每个标签会通过两个子载波频率中的一个来发送一位。一个载波频率专门用来发送 1，另一个载波频率专门用来发送 0，以便于阅读器能够同时识别 0 和 1。如果阅读器同时接收到 0 和 1，阅读器会把确认位 0 发送给标签，否则，阅读器会用接收到的位的值来代替原有的确认位。只有第一位与确认位匹配的标签才能响应，并把接下来的位发送给阅读器，其他标签阅读器会暂时进入沉默状态，直到阅读器在一个新的标签识别循环中，请求标签重新开始响应为止。上述的过程将会一直重复，直到其中的一个标签 ID 号的所有位被成功识别为止。然后此标签将进入休眠状态，直到阅读器发送请求信号，使所有的标签重新开始，进入下一个识别过程。

接下来，将给出一个例子，来解释 EPCglobal Class 0 的细节。假设，有三个标签，ID 号分别为 001、011 和 110。下面将描述标签的识别过程。

1）一开始，阅读器会发送一个请求信号给标签，以开始一个标签的识别循环。接收到请求后，标签会把它 ID 号的第一位发送给阅读器。具体地说，标签 1（ID 号为 001）会发送回“0”，标签 2（ID 号为 011）会发送回“0”，标签 3（ID 号为 110）会发送回“1”给阅读器。

2）阅读器会从两个独立的子载波信道中同时接收到“0”和“1”，并把确认位“0”发送给标签。标签 1 和标签 2 会把它们 ID 号的第二位发送给阅读器（标签 1 发送回“0”，标签 2 发送回“1”）。标签 3 则进入沉默状态，直到接收到下一个请求信号为止。

3）阅读器仍旧同时接收到“0”和“1”，并再次把确认位“0”发送给标签。然后，标签 1 会把它 ID 号的第三位发送给阅读器，而标签 2 进入沉默状态。

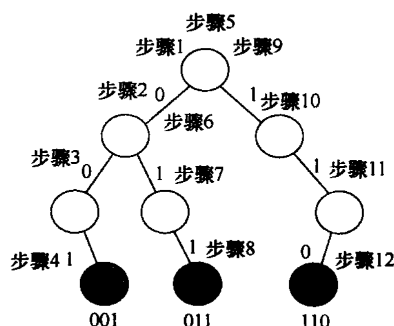
4）由于只有标签 1 发送回“1”，并且响应位的长度与 ID 号的长度相同，阅读器识别到“1”，并把标签 1 的 ID 号（001）设置成已识别状态。当收到阅读器的确认位，并发送它 ID 号的最后一位之后，直到下一个识别过程开始为止，标签 1 都将进入休眠状态。

5）阅读器请求标签开始新一轮的识别循环。所有沉默中的标签都将会开始响应阅读器。

识别过程的步骤将会持续，直到识别区域内所有的标签都被成功地识别为止。表 2-3 为整个识别过程的步骤和相关的分裂树图。

表 2-3　EPCglobal Class 0 协议的识别过程

步骤	命令/确认比特	响应比特	状态
1	新一轮	标签 001:0 标签 011:0 标签 110:1	碰撞
2	0	标签 001:0 标签 011:1 标签 110:哑的	碰撞
3	0	标签 001:1 标签 011:哑的 标签 110:哑的	成功
4	1	标签 001:睡眠的 标签 011:哑的 标签 110:哑的	识别 (标签 001)
5	新一轮	标签 001:睡眠的 标签 011:0 标签 110:1	碰撞
6	0	标签 001:睡眠的 标签 011:1 标签 110:哑的	成功
7	1	标签 001:睡眠的 标签 011:1 标签 110:哑的	成功
8	1	标签 001:睡眠的 标签 011:睡眠的 标签 110:哑的	识别 (标签 011)
9	新一轮	标签 001:睡眠的 标签 011:睡眠的 标签 110:1	成功
10	1	标签 001:睡眠的 标签 011:睡眠的 标签 110:1	成功
11	1	标签 001:睡眠的 标签 011:睡眠的 标签 110:0	成功
12	0	标签 001:睡眠的 标签 011:睡眠的 标签 110:睡眠的	识别 (标签 110)



2.4.2.4　TSA 协议

树时隙 ALOHA（TSA）协议[17]是一种混合的协议，它集成了树分裂和动态帧时隙 ALOHA 协议的原理。在 TSA 中，阅读器首先会初始化帧的 S 的大小（S 表示帧中时隙的数量），把它发送给所有的标签，并请求它们的 ID 号。在接收到请求后，所有的标签会随机地选择一个标记为 1 ~ S 的时隙来发送它们的 ID 号。如果在一个时隙中，只有一个标签响应，那么标签将会被成功地识别。但是，如果有多个

标签在一个时隙时，阅读器会记住这个时隙号，并命令这些标签在下一个帧中响应。需要注意的是，帧的大小是通过参考文献［14］定义的一个专门的估计函数来计算的。如果在这个帧中仍然存在着标签的碰撞，同样的动作将会进行，来分裂碰撞的标签。这个动作与根据一个树结构来把碰撞的标签分裂成多个子群的方法相似。这就是为什么称它为树时隙 ALOHA 的原因。

在 TSA 协议中，阅读器包含帧大小的请求、分裂碰撞标签的时隙数和标签分裂树的各层。通过存储选择的时隙数和保持一个标签分裂树的可变层，标签能够跟踪识别过程的状态。因此，识别的过程能够正常地进行，所有的标签也都将能被成功地识别。

2.4.2.5 BSQTA 和 BSCTTA 协议

BSQTA 和 BSCTTA 协议是 Choi 等人在参考文献［18］中提出的，目的是为了改善 QT 协议的性能。在 QT 协议的识别过程中，当阅读器发送长度为 k 的请求字符串 S 给标签时，ID 前缀与 S 匹配的标签将会传回它的位 $k+1$，…，n 的部分 ID 号给阅读器，其中 n 是 ID 号的长度。如果碰撞发生，阅读器会发送请求位字符串 S0 和 S1 给标签。Choi 等人在参考文献［20］中的研究表明，请求位字符串 S0 和 S1 与前 k 位相同，而与最后一位不同。基于这个发现提出了 BSQTA 和 BSCTTA 协议，通过使用两个响应的时隙，来减少识别的时间。接下来，将介绍这两个方法的识别过程。

1）阅读器发送长度为 b-1 的位字符串 S 给标签。

2）在阅读器识别区域内的标签，如果它的 ID 号的前 b-1 位与 S 匹配，则它会在两个时隙中选择一个时隙把 ID 号发送给阅读器。如果 ID 的第 k 位为 0，标签会在第一个时隙内响应，否则，它会在第二个时隙内响应。

① 对于 BSQTA 来说，标签会发送 ID 号的第 $b+1$ 到最后一位。

② 对于 BSCTTA 来说，标签会发送 ID 号的第 $b+1$ 到最后一位，直到此标签收到一个确认命令才停止发送。这个确认命令由阅读器发送，表明发生了碰撞。

3）如果在时隙内没有碰撞，标签将会被成功地识别。

4）如果在标记为 0 或 1 的时隙内碰撞发生，阅读器随后将会发送一个新的请求位字符串给标签。

① 对于 BSQTA 来说，新的请求位字符串会在 S 的基础上再加上这个时隙号（0 或 1）。

② 对于 BSCTTA 来说，新的请求位字符串会在 S 的基础上再加上碰撞之前已经接收的位字符串。

上述过程会一直重复，直到所有的标签都被成功地识别为止。就像参考文献［25］中指出的，BSQTA 和 BSCTTA 协议能够提高 QT 协议的性能。

2.4.2.6 AQS 协议

AQS 协议是由 Myung 等人在参考文献［24］中提出的一种具有适应性的标签

防碰撞协议。这个协议的基本原理是参考从最后一次识别循环中获得的标签信息来减少碰撞，这个协议的前提是假设在一连串循环中，标签的数量不发生显著的变化。AQS 协议的识别过程与 QT 协议相似，主要的区别在于待发送队列中的请求位字符串是从最后一个识别循环中复制过来的。队列不仅包括成功识别标签的各步骤的请求位字符串，也包括了没有任何标签响应各步的请求位字符串。如果识别区内标签的数量保持不变，那么不必修改队列中的任何请求位字符串，所有的标签都将被成功地识别。但是在最后一轮循环后，如果有标签进入或是离开识别区，那么就必须执行下面所列的动作。

1）有标签进入识别区时：如果使用最后一轮识别循环提供的请求位字符串 S 发送碰撞，那么可以确定，在最后一轮识别循环后，有新的标签进入了阅读器的识别区。在这种情况下，树分裂过程将会进行，并且更长的请求字符串也将会加入到队列中。

2）有标签离开识别区时：如果有标签离开识别区，最后一轮识别循环提供的位字符串 S 的一些请求将不会被响应。为了提高识别的性能，阅读器需要把 S 与队列中的另一个字符串合并，这个字符串除了最后一位，其他各位都与 S 相同。

2.4.3　基于计数器的协议

基于计数器的协议[11,22-24]与基于树的协议相似，不存在标签饥饿的问题。此协议的两类基本的思想是把遇到碰撞的标签分裂成多个子群，直到在一个子群中，只到其中 1 个标签能够成功地被识别为止。这两类协议主要的不同是，基于树的协议是使用静态的标签 ID 号来进行确认性地分裂，而基于计数器的协议是使用动态改变的计数器来进行概率性地分裂的。由于基于计数器的协议不需要使用标签 ID 号来进行分裂，所以它具有稳定的性能。这个性能不会被标签 ID 号的分布和 ID 号的长度影响。本节将介绍两个基于计数器的标签防碰撞协议：ISO/IEC 18000-6B 和 ABS 协议。

2.4.3.1　ISO/IEC 18000-6B 协议

ISO/IEC 18000-6B 协议[22]采用了基于计数器的标签防碰撞协议。在 ISO/IEC 18000-6B 协议中，每个标签使用动态改变的计数器和一个随机位产生器来识别标签。所有标签的计数器都初始化设置为 0，并且计数器值 0 的标签能够发送它的 ID 号给请求的阅读器。当碰撞发生时，阅读器会关注所有碰撞的标签。计数器值大于 0 的标签随后会使它们的计数器值加 1，而计数器值为 0 的标签会随机地产生一个随机位 0 或 1，并把此位加到计数器上。在这种方式下，计数器值为 0 的标签会分裂成两个子群，一个子群的标签计数器值为 0，另外一个子群的标签计数器值为 1。分裂的过程将会持续的进行，直到计数器值 0 的标签只有一个，或是一个也没有为止。在只有一个标签的计数器值为 0 的情况时，这个标签能够被成功地识别，并在随后保持沉默，直到标签的识别过程结束为止。计数器值为 0 只有一个标签或是一

个标签也没有的情况中，阅读器会发送一个命令，通知所有的标签，使它们的计数器值都减1。这个过程会一直持续，直到所有的标签都被成功地识别为止。

接下来，将用一个例子来说明ISO/IEC 18000-6B协议的整个过程。假设，有四个ID号为0010、0110、1001和1110的标签。标签的识别过程描述如下。

1）一开始，阅读器会发生一个请求，使所有的标签开始一个标签识别循环。在接收到请求后，标签会把它的计数器值设置为0。标签1（ID号为0010）、标签2（ID号为0110）、标签3（ID号为1001）和标签4（ID号为1110）会同时响应，并把它们的ID号发送给阅读器。这时会发生碰撞。

2）阅读器会发送一个碰撞通知命令给所有的标签，并使标签的计数器值随机地加0或加1。标签1、标签2、标签4的计数器值0，它们会同时响应ID号给阅读器，随之，碰撞会再次发生。

3）阅读器会发生一个碰撞通知命令给标签1、标签2和标签4，随机地使它们的计数器值增加0或1。由于标签3的计数器值已经增加1，计数器值为0的标签2会把它的ID号发送给阅读器，然后它将会被成功地识别。

4）阅读器会把一个成功识别的确认命令发送给已经成功识别的标签2。随后，标签2将会进入沉默状态，其他未识别的标签1、标签3和标签4会把它们的计数器值减1。

识别过程将会重复，直到所有的标签都被成功地识别为止。表2-4为识别过程的整个步骤和相关的树图。

表2-4 ISO/IEC 18000－6B的识别过程

步骤	阅读器命令	标签ID	计数器值	随机比特	新的计数器值	响应	树图
1	请求	1	-		0	0010	0 0010 0110 1001 1110
		2	-		0	0110	
		3	-		0	1001	
		4	-		0	1110	
2	碰撞	1	0	0	0	0010	0 0010 0110 1110 1 1001
		2	0	0	0	0110	
		3	0	1	1		
		4	0	0	0	1110	
3	碰撞	1	0	1	1		2 1001 0110 0 1 0010 1110
		2	0	0	0	0110	
		3	1		2		
		4	0	1	1		

（续）

步骤	阅读器命令	标签 ID	计数器值	随机比特	新的计数器值	响应	树图
4	成功	1 2 3 4	1 0 2 1		0 – 1 0	0010 1110	1 1001 0 0010 1110
5	碰撞	1 2 3 4	0 – 1 0	0 1	0 – 2 1	0010	2 1001 0 0010　1 1110
6	成功	1 2 3 4	0 – 2 1		– – 1 0	 1110	1 1001 0 1110
7	成功	1 2 3 4	– – 1 0		– – 0 –	 1001	0 1001
8	成功	1 2 3 4	– – – –				

2.4.3.2　ABS 协议

ABS 协议[24]的提出是用来提高 ISO/IEC 18000-6B 标签防碰撞协议的性能。在 ABS 协议中，标签会使用两个计数器，分别是优化时隙计数器（PSC）和分配时隙计数器（ASC）。PSC 表示已经成功识别的标签数量，初始值为 0，当一个标签被成功识别后，PSC 会加 1。使用 PSC 和 ASC 之后，标签可以决定它是否能够发送它的 ID 号给阅读器。当标签的 ASC 与 PSC 相等时，它就能够发送它的 ID 号。如果没有

一个标签响应，那么所有ASC大于PSC的标签将会使ASC减1。当碰撞发生时，阅读器会通知所有碰撞的标签。这种情况下，所有ASC大于PSC的标签随后会使ASC加1，而ASC与PSC相等的标签，则会使ASC随机加0或者1。需要注意的是，ASC比PSC小的标签在识别过程完成之前不会使ASC增加，甚至不会发送自己的ID号，因为这些标签之前已经被识别了。

所有的标签被成功识别后，标签之间将会有唯一的、连续的ASC值。这些ASC值将会保留在下一个标签识别循环中使用，以加快标签的识别过程。如果在最后次识别过程后，有标签进入或是离开识别区域，下面的动作将必须执行。

1）当有新标签进入识别区时：当有一个新的标签进入识别区，并且接收阅读器的请求后开始新的识别循环时，它会把它的PSC设置为0，并把它的ASC设置成一个随机的值R（R的范围由阅读器提供）。此新的标签会和ASC的值为R的旧标签发生碰撞。上述提及的ABS协议过程通过调节所有标签计数器的值来解决碰撞。

2）当有标签离开识别区时：如果阅读器检测到没有标签响应时，阅读器就可以确定是否有标签离开了识别区。ASC大于PSC的所有标签将会使它的ASC值减少1，来解决这种状况。

参考文献［24］报道，ABS协议可以显著地提高ISO/IEC 18000-6B标签防碰撞协议的性能。这验证了，在连续的识别循环中，当标签的数量不是变化得非常大的时候，最后一次识别循环的计数器信息是非常有用的。

2.5 结论

怎样在识别区内快速准确地识别标签是阅读器和标签防碰撞协议设计的基础。为了满足这个需要，已经提出了许多防碰撞协议。本章已经介绍了多个基于TDMA、FDMA和CSMA的阅读器防碰撞协议，也介绍了一些基于ALOHA、树和计数器的标签防碰撞协议。表2-5列出了本章提及的各种协议的特性。接下来，将给出一个总结，并指出这些协议的新的研究方向。

表2-5 本章介绍的协议的特性

协议类型	协 议	特 性
阅读器防碰撞	DCS[7] Colorwave[7]	在Colorwave（与DCS相比）中，颜色的数量能够动态调整（DCS不能） 协议假设阅读器之间是时间同步的 阅读器很难单独检测到碰撞
	HiQ[8]	应用到移动环境中时，效率很低 由于分层结构，付出更多的管理代价
	EPCGlobal Gen 2[9]	不适用于低成本的没有频率选择功能的标签
	ETSI 302 208 Standard[10]	阅读器很难仅仅依靠感知载波来检测碰撞

（续）

协议类型	协　议	特　性
标签防碰撞	ALOHA[12] 时隙 ALOHA[13] 帧时隙 ALOHA[14] ISO/IEC 18000-6A[22]	算法简单 会出现标签饥饿问题 时隙 ALOHA 的执行性能比 ALOHA 好两倍 帧时隙 ALOHA 需要恰当地评估帧的大小来获得更好的执行性能
	QT[16]	算法和标签电路简单 不需要非易失性可写标签内存 识别延迟收到分发和标签 ID 长度的影响
	逐步二进制树[15] EPCglobal Class 0[19]	更低的消息通信 复杂度依赖于标签 ID 的长度而不是它的分发
	TSA[17]	标签需要一个复杂的电路
	BSQTA BSCTTA[18]	标签需要一个复杂的电路 需要时间同步 在 BSCTTA 中，标签很难同时监听信道和进行通信
	AQS ABS[30]	如果标签数量不会改变太多，协议的执行效果非常好
	ISO/IEC 18000-6B[22]	执行性能不会受到标签 ID 长度和分发的影响 不会出现标签饥饿问题

2.5.1　阅读器防碰撞协议的总结和新的研究方向

由于无源标签尺寸较小，并且是由阅读器来提供能量的，所有它只具有有限的计算和通信能力。通常在无线通信领域用来避免碰撞的 RTS/CTS 的机制不适合 RFID 系统。RFID 系统需要有新的机制来减少碰撞。2.3 节介绍了一些减少阅读器碰撞的阅读器防碰撞协议。它们可以分为 TDMA、FDMA 和 CSMA 协议。随着 RFID 的广泛使用，阅读器防碰撞协议也有了新的研究方向。

1. 移动环境下的阅读器防碰撞协议

阅读器在静止的环境中，是可以分配例如频率信道和时隙这样的资源来尽可能地减少碰撞。但是，如果阅读器移动的话，信号产生的干扰将会是动态的，并且是不可预测的。之前提出的方案都不适用于阅读器在动态的环境下，因而需要有新的阅读器防碰撞协议应用于此环境下。

2. 高密度环境下的阅读器防碰撞协议

在高密度的环境中，怎样公平地在阅读器之间分配，例如频率信道、时隙这样的资源是非常复杂的。在这样的环境中，需要更加有效的阅读器防碰撞协议。在某些情况中，一个阅读器需要与其他阅读器相互协作来跟踪标签，它们之间的协作可以扩大标签被跟踪的范围。

3. 具有手持式设备和有源标签环境下的阅读器防碰撞协议

手持式设备的一个最大的不足是它的能力受限。因此在具有手持式设备的环境中，需要具有能量效率和能量意识的阅读器防碰撞协议来节省手持式设备的能量。节能和能量意识也可以使用到有源标签的环境中，以提高它们的使用时间。

2.5.2 标签防碰撞协议的总结与新的研究方向

2.4节把标签防碰撞协议分为基于ALOHA、基于树和基于计数器这三类。基于ALOHA的协议较简单，且具有平衡的性能，但是，它们存在标签饥饿问题，即如果某个标签的响应总是与其他标签碰撞，将永远不会被识别。与基于ALOHA的协议相比，基于树的协议需要更长的识别时间，但是基于树的协议不存在标签饥饿问题。基于树的协议存在的不足是它们的性能受到标签ID号的长度和分布的影响。与基于树的协议相似，基于计数器的协议也不存在标签饥饿问题，同时具有稳定的性能，因为它们并不受标签ID号的长度和分布的影响。表2-6根据时间复杂性、信息复杂性和标签是否需要可写能力的存储器，比较了一些典型的标签防碰撞协议。表格内相关的数据引自参考文献［14，18，20，31-35］。

表2-6 标签防碰撞协议关于时间复杂度、消息复杂度和标签的非易失可写内存需求的比较

协议	时间复杂度	消息复杂度	对标签非易失可写内存的需求
QT[16]	平均情况：$O(n)$ 最坏情况： $n\times(k+2-\log n)$[31,32]	最坏情况： $k\times(2.21\log n+4.19)$[31,32]	不需要[33]
逐步二进制树[15]	最坏情况： $\theta(2^k)$[32,34]	平均情况： $O(n(k+1))$[32,34]	需要（8bit内存作为指针，来维护最后的比特发送[18]）
ISO/IEC 18000-6B[22]	$O(n)$[20,32]	最坏情况：[32,35] 当n未知时，为θ（$n\log n$） 当n已知时，为θ（n）	需要（8bit或16bit内存作为计数器来保存标签响应顺序和计时[30]）
帧时隙[14]	上限： $t\times s$ + time required 来评估N[14,32]	上限： $n\times s$[32]	需要（8bit或16bit内存作为计数器来保存标签响应的时隙[14]）

注：n是识别区中标签的数量，k是标签ID的长度，t是1帧的时长，s是识别所有标签所需的帧数量，N是n的估计值。

一个优秀的标签防碰撞协议需要某些特性。在研究一个新的标签防碰撞协议时，需要牢记如下特性。

1）阅读器需要识别所有在它识别区域内的标签。由于标签饥饿问题这样的原因，某些标签不能成功地识别，在一些应用中会造成问题。因此，一个优秀的标签防碰撞协议不会落下对任何一个标签的识别。

2）在许多应用中，标签经常是附着在移动的物体上的。由于阅读器只能够成功地识别在它的识别区域内的标签，所以阅读器需要能够尽可能快地识别标签，以便于在移动标签离开阅读器识别区域之前，能够被成功地识别到。

3）由于标签通信和计算能力的限制，标签防碰撞机制不能过于复杂。就是说，在设计时，必须使标签防碰撞协议尽量地简单。

4）当一个不怀好意的人使用阅读器识别到一个附着标签的物品时，这个物品拥有者的隐私将会受到侵害。因此，需要具有隐私保护机制的防碰撞协议，以便于在不泄露隐私的情况下，标签能够有效地被识别。

5）如果在连续的标签识别过程中，标签的数量没有显著的改变，最后一次识别循环获得的信息将非常有用。可以使用这些信息，设计出具有更快识别过程的协议。但是，在设计这类协议时，必须考虑到标签进入和离开识别区的情况。

参考文献

1. P. Schaar. Working document on data protection issues related to RFID technology. In *Working Document Article 29—10107/05/EN*. European Union Data Protection Working Party, January 2005.
2. H. Chae, D. Yeager, J. Smith, and K. Fu. Maximalist cryptography and computation on the WISP UHF RFID tag. In *Proc. of the Conference on RFID Security*, Málaga, Spain, 2007.
3. S. M. Birari and S. Iyer. PULSE: A MAC protocol for RFID networks. In *Proc. of the EUC Workshops*, pp. 1036–1046, Nagasaki, Japan, 2005.
4. D. Y. Kim, B. J. Jang, H. G. Yoon, J. S. Park, and J. G. Yook. Effects of reader interference on the RFID interrogation range. In *Proc. the 37th European Microwave Conference (EuMC'07)*, pp. 728–731, Munich, Germany, 2007.
5. D. Y. Kim, H. G. Yoon, B. J. Jang, and J. G. Yook. Interference analysis of UHF RFID systems. *Progress in Electromagnetics Research*, 4:115–126, 2008.
6. Y. Tanaka and I. Sasase. Interference avoidance algorithms for passive RFID systems using contention-based transmit abortion. *IEICE Transactions*, 90-B(11):3170–3180, 2007.
7. J. Waldrop, D. W. Engels, and S. E. Sarma. Colorwave: An anti-collision algorithm for the reader collision problem. In *Proc. of IEEE International Conference on Communications*, Vol. 2, pp. 1206–1210, Anchorage, Alaska, 2003.
8. S. E. Sarma, J. Ho, and D. W. Engels. HiQ: A hierarchical Q-learning algorithm to solve the reader collision problem. In *Proc. of SAINT Workshops*, pp. 88–91, Phoenix, AZ, 2006.
9. EPCglobal. *EPCglobal Class-1 Generation-2 UHF RFID Protocol*, April 2004. Version 1.0.9.
10. ETSI. *EN 302 208-2 Protocol*, September 2004. Version 1.1.1.
11. M.-K. Yeh and J.-R. Jiang. Adaptive *k*-way splitting and pre-signaling for RFID tag anti-collision. In *Proc. of the 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON'07)*, Taipei, Taiwan, 2007.

12. N. Abramson. The ALOHA system-another alternative for computer communications. In *Proc. of Fall Joint Computer Conference of AFIPS*, Vol. 37, pp. 281–285, Houston, TX, 1970.
13. L. Liu and S. Lai. ALOHA-based anti-collision algorithms used in RFID system. In *Proc. of Int'l Conf. on Wireless Communications, Networking and Mobile Computing 2006 (WiCOM 2006)*, pp. 1–4, Wuhan, China, 2006.
14. H. Vogt. Efficient object identification with passive RFID tags. In *Proc. of Pervasive Computing*, pp. 98–113, Berlin, 2002.
15. H. Choi, J. R. Cha, and J. H. Kim. Fast wireless anti-collision algorithm in ubiquitous ID system. In *Proc. of IEEE VTC*, Los Angeles, CA, 2004.
16. F. Zhou et al. Evaluating and optimizing power consumption of anti-collision protocols for applications in RFID systems. In *Proc. of the 2004 International Symposium on Low Power Electronics and Design*, New York, 2004.
17. M. A. Bonuccelli, F. Lonetti, and F. Martelli. Tree slotted Aloha: A new protocol for tag identification in RFID networks. In *Proc. of the 4th IEEE International Workshop on Mobile Distributed Computing (MDC'06)*, New York, 2006.
18. J. H. Choi, D. Lee, and H. Lee. Bi-slotted tree based anti-collision protocols for fast tag identification in RFID systems. *IEEE Communications Letters*, 10(12):861–863, 2006.
19. Draft protocol specification for a 900 MHz class 0 radio frequency identification tag, Auto-ID Center, Cambridge, MA, Technical report, 2003.
20. D. H. Shih, P. L. Sun, and D. C. Yen. Taxonomy and survey of RFID anti-collision protocols. *Computer Communications*, 29(11):2150–2166, 2006.
21. D. Krebs and M. J. Liard. *White Paper: Global Markets and Applications for Radio Frequency Identification*. Venture Development Corporation, 2001.
22. ISO/IEC. Information technology automatic identification and data capture techniques—radio frequency identification for item management air interface—part 6: parameters for air interface communications at 860–960 MHz. Final Draft International Standard ISO 18000-6, November 2006.
23. *Philips Semiconductors, UCODE*, http: // www. semiconductors. philips. com, 2005.
24. J. Myung and W. Lee. Adaptive splitting protocols for RFID tag collision arbitration. In *Proc. of MobiHoc 2006*, pp. 202–213, Florence, Italy, 2006.
25. *Dense RFID Reader Deployment in Europe using Synchronization*, January 2008. Final draft ETSI EN 302 208-1 V1.2.1.
26. J. R. Cha and J. H. Kim. Novel anti-collision algorithms for fast object identification in RFID system. In *Proc. of the 11th International Conference on Parallel and Distributed Systems—Workshops (ICPADS'05)*, pp. 63–67, Fuduoka, Japan, 2005.
27. G. Khandelwal et al. ASAP: A MAC protocol for dense and time constrained RFID systems. In *Proc. of IEEE International Conference on Communications (ICC'06)*, Istanbul, Turkey, 2006.
28. S. Lee, S. D. Joo, and C. W. Lee. An enhanced dynamic framed slotted aloha algorithm for RFID tag identification. In *Proc. of Mobiquitous*, pp. 166–172, 2005.
29. L. G. Roberts. Extensions of packet communication technology to a hand held personal terminal. In *Proc. of AFIPS Spring Joint Computer Conf.*, Vol. 40, pp. 295–298, Montvale, NJ, 1972.

30. M. Kodialam and T. Nandagopal. Fast and reliable estimation schemes in RFID systems. In *Proc. of ACM Mobicom*, Los Angeles, CA, 2006.
31. C. Law, K. Lee, and K. Y. Siu. Efficient memoryless protocol for tag identification. In *Proc. of the 4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communication*, pp. 75–84, Boston, MA, August 2000.
32. C. Abraham, V. Ahuja, A. K. Ghosh, and P. Pakanati. Inventory management using passive RFID tags: A survey. Technical report, Technical Report of Department of Computer Science, The University of Texas at Dallas, Dallas, TX, 2003.
33. K. Ali, H. Hassanein, and A. Taha. RFID anti-collision protocol for dense passive tag environments. In *Proc. of 32nd IEEE Conference on Local Computer Networks*, Dublin, Ireland, 2007.
34. M. Jacomet, A. Ehrsam, and U. Gehrig. Contact-less identification device with anti-collision algorithm. In *Proc. of IEEE International Conference on Circuits, Systems, Computers and Communications*, Sado Island, Niigata, July 1999.
35. D. Hush and C. Wood. Analysis of tree algorithms for rfid arbitration. In *Proc. of IEEE International Symposium on Information Theory*, Cambridge, MA, August 1998.

第3章　用于RFID的低功耗转发器

由于无线射频识别（RFID）技术是一种廉价的可靠的自动化识别解决方案，它在制造业、仓储业和物流业中广受欢迎。此外，RFID转发器在普适计算中，也将是一个重要的设备。标签可以用来采集感知数据，对非侵害性环境的监视。由于低功耗无源超高频（Ultra-High Frequency，UHF）转发器具有更大的识别范围，它在上述应用中将扮演重要的角色。对于无源标签来说，它的能量是在通信期间，从阅读器辐射的电磁波中获得的，这些能量需要供给数字电路和其他感知设备使用。因此，需要实现有效的低功耗策略，来提高通信范围和电路的功能。本章测试转发器设计的细节，深入介绍一个低功率、工作在超高频带、服从标准协议（ISO 18000-6B和-6C协议）的标签来支持设备协同工作。最后提出一些在高执行性能无处不在的感知转发器中的开放性研究问题。

3.1　概述

RFID技术已经在制造业、物流、运输等领域开始逐渐流行，因为这些领域需要对物品进行精确地识别和跟踪[1]。RFID标签对于全自动供应链来说，是一个低成本的解决方案，它可以克服读写距离等的限制，从而成为替代条形码技术的一个候选方案[2]。不过，RFID技术除了替代条形码之外，还有着其他的功能。普适计算的目的是创造一个“动态的环境”，从而将集成计算设备分布至现实环境中。RFID技术将在非侵害性环境监测中扮演重要的角色：标签可以嵌入具有计算能力、数据存储功能的传感器和通信设备中使标签具有“智能”，并且是一个廉价的网络节点。特别是当转发器工作在超高频（UHF）频带时，读写的范围将扩大到数米，从而提高了它的识别能力。为了在民用化市场渗透，每个标签的成本应尽量地降低，为了满足这个苛刻的限制，制造过程必须廉价。在大部分应用中，由于有源标签包含价格较高的电池，使得只能选择无源标签。高性能的无源标签由一个很小的集成电路芯片、灵活的印制天线和一个为应用设计的底片[1]组成。标签电路工作时所需要的能量，是通信期间从阅读器辐射的电磁波中获得的。因此，需要实现一个有效的低功耗运行策略，来提高通信范围和电路的功能。虽然一个标签包括了一个模拟的射频前端和一个存储器，但是能量需求最大的部分，往往是基带处理器[2]。因此，设计一个超低功耗的基带处理器给感知部分提供了更高的能量，可以提高测量的执行性能。

本章展示转发器的最新设计。首先，对系统进行深入的分析，随后介绍转发器

设备的特性和设计的目标。接下来，介绍基于低电压运行和低电流设计技术的电路设计方案。讨论包括传感器集成和安全支持在内的开放的研究问题和一些潜在的解决方案。

3.2 关于最新的RFID实现的调查

现在，已经有众多运行在工业、科学和医疗领域（ISM）的13.56MHz[1,3]的RFID系统。这些系统在标签和阅读器之间依靠电容或是电磁耦合，来实现近距离的通信。虽然高频（High Frequency，HF）系统对于许多应用来说也是可靠和合适的。但是耦合方式实质上限制了它们的读写最大范围只能在1m左右。现在应用于有源识别的系统，使用的是例如868/915MHz[4-13]和2.4GHz[14,15]的ISM频带这样的更高的频率。超高频（UHF）系统允许阅读器与标签之间在更远的范围内使用电磁波来进行耦合，因此可以实现更远距离的通信。此外，还可以使用更大范围的带宽，这可以使通信获得更高的数据率。

通过标签和阅读器间使用私有的通信协议，人们已经在完整的RFID芯片上做了相当多的工作。Kocer等人[4]报道了无线的遥感勘测设备，它可以从一个450MHz的射频信号中获得能量和一个参考时钟，并能利用900MHz的二进制相位键控（Phase Shift Keying，PSK）调制的载波来获得一个ID数据。Karthaus等人[5]详细报道了一个工作于超高频的低功耗RFID标签，这个标签具有EEPROM存储器和依靠在前向链路（阅读器到标签的通信）上使用脉宽调制（Pulsewitlth Modulation，PWM），在反向链路（标签到阅读器的通信）上使用PSK调制的完整的通信协议。Curty等人[15]详细介绍了一个工作于2.4GHz的远端供能的，在前向链路上使用开关键控（On-Off Keying，OOK）调制，反向链路上使用振幅键控（Amplitude，Shift Keying，ASK）调制的RFID转发器。在RFID的应用[6-8]中，人们已经做了相当多的基于CMOS模拟前端的工作。

对已发布的和广泛使用的标准以及支持设备间的互操作的关注，是将来工作的基础。Gillen等人[2]报道了一个较好的RFID芯片，它可以实现EPC Class 0协议[16]，而参考文献［10-12］则是集中关注对完整的数字基带处理器的优化。参考文献［9］报道了一种与ISO 18000-6B兼容的，具有超低功耗电池/无源标签，它能够同时工作于高频和微波频带。

3.3 RFID系统需求

RFID标签主要的设计目标总结如下。

1）宽的读范围，即标签能够从天线范围内有效地获得能量。

2）低成本，即与超大规模集成电路（VLSI）低成本制造技术相兼容，并具有

小尺寸设备的引脚。

3）设备的互操作性允许在阅读器网络内通信。

这些需求将会在接下来详细说明。

3.3.1 电磁传播基础和标签能量消耗

一个电磁（EM）波从它的源点向空间传输的方式是球形的。天线均匀地向各向辐射能量大小为 P_T 的称为全向或是球形的电磁波。距离此发射源距离为 r 的任意点，辐能流率（曾称为辐射密度）可以表示为

$$S = \frac{P_T}{4\pi r^2} \tag{3-1}$$

这种情况称为远场，因为 r 比电磁波波长 λ 的两倍还要大。相反的，在近场中，电磁波的电场和磁场元素的关系和辐能流率将很难确定。首先假设是分析在远场下的情况。

实际上，不存在物理上的全向天线：因为每一个发射源都有它的辐射模式，任何方向都有一定的增益 $G_T(\theta,\phi)$，这个增益用角 θ 和 ϕ 表示。天线的方向性由它的物理特性决定。因此，沿着主要的辐射方向，辐能流率是全向发射源的 G_T 倍，即

$$S = \frac{P_T G_T}{4\pi r^2} \tag{3-2}$$

有效的全向辐射源，简称为EIRP，能够用能量来定义。这个能量由全向天线提供来感知由辐射源提供的在距离为 r 处的同样的辐能流率：

$$P_{T,EIRP} = P_T G_T \tag{3-3}$$

例如一个偶极天线，它的最大增益 $G_T = 1.64$，因此，它的 $P_{T,EIRP} = 1.64 P_T$。实际上，偶极天线发射的能量也被称为有效的辐射能量（ERP），因此 $P_{T,EIRP} = 1.64 P_{T,ERP}$。另外一方面，一个匹配状况下的接收天线探测到辐能流率 S，并把有效的能量传递给它的负载[18]，即

$$P_{AV} = S \frac{\lambda^2}{4\pi} G_{ANT} \tag{3-4}$$

式中 λ——接收的电磁波的波长；

G_{ANT}——接收天线的增益（假设天线是朝最大增益的方向）。

结合式（3-2）和式（3-4）可以得出：

$$P_{AV} = P_{T,EIRP} \left(\frac{\lambda}{4\pi r}\right)^2 G_{ANT} \tag{3-5}$$

式（3-5）被称为Friis传输方程[19]，描述了由波长为 λ，能量为 $P_{T,EIRP}$ 的电磁波发射的，在距离发射源为 r 处的一个增益为 G_{ANT} 的天线所接收到的能量大小。阅读器发射由一定载波构成的电磁波，这个电磁波的频率被调制在860～960MHz

之间，能量在500mW ERP和4W EIRP之间（这根据本地的规定）[20]。假设转发器具有一个统一增益的天线（G_{ANT}），它的可用能量［由式（3-5）给出］由图3-1指出的读距离表示。需要注意的是，当距离大于数米时，这个能量会迅速下降到μW的范围。因此，无源标签的能量感应效果必须要满足最后成品对于识读范围的要求。在这个结构中，使标签核心电路工作在稍高于阈值电压，这可以在保持均衡性能[21]的基础上减少能量的消耗。

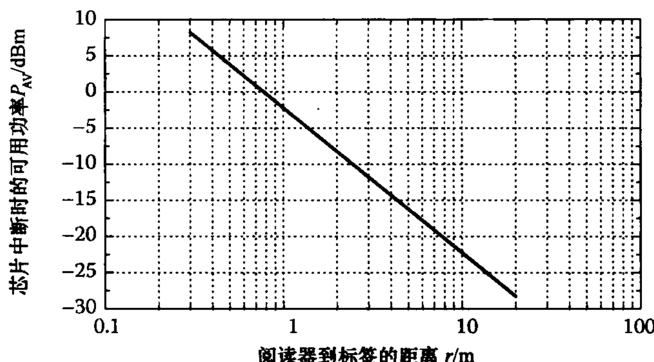


图3-1 一个与阅读距离相比有500mW EPR传输功率的0dB匹配天线的可用功率

图3-2描述了一种良好的无源转发器框图。如上述已经强调的，提供给数字电路的能量，是由对天线端（V_{ANT}）收集的电压进行校正和转换来的。一个片上倾销电容（C_{DUMP}）用来作为能量存储元件，以便在标签识别和响应期间保持足够的能量供给。倾销电容必须在一定时间间隔内是可以充电的，以便于具有一定的电压在识别开始之前（$t = t_0$）给RFID上电。这个初始状态如式（3-6）所示，即

$$V_{DD,t_0}^{unreg} = V_{DD,MAX}^{unreg} = \frac{Q_{MAX}}{C_{DUMP}} \tag{3-6}$$

此外，假设使用的是一个线性的电压调节器，具有一个最小的输入电压，

$$V_{DD}^{unreg}(t) > V_{DD,MIN}^{unreg} \tag{3-7}$$

在运行期间一直维持，来保证产生正确的电压值（V_{DD}^{reg}），理论上说，标签可用的瞬间能量（$dE_{OUT}(t)/dt$）是有限的，且受到输入能量（$dE_{IN}(it)/dt$）的限制，而标签从阅读器发送辐射的能量中获得的总能量是无限的。倾销电容扮演着能量储存器的角色，可以增加标签电路的瞬间能量，因此面临一个最高能量请求，它可以容忍校正电压（V_{RECT}）大于最小值$V_{RECT,MIN}$。这种状况是当标签的能量消耗$E_{OUT}(t)$被限制在式（3-8）时实现的：

$$E_{OUT}(t) < E_{IN}(t) + \frac{C_{DUMP}}{2}((V_{DD,t_0}^{unreg})^2 - (V_{DD,MIN}^{unreg})^2) \tag{3-8}$$

通过整个读过程（也就是，当$t_0 < t < t_{END}$）。式（3-8），需要的能量$E_{OUT}(t)$能够

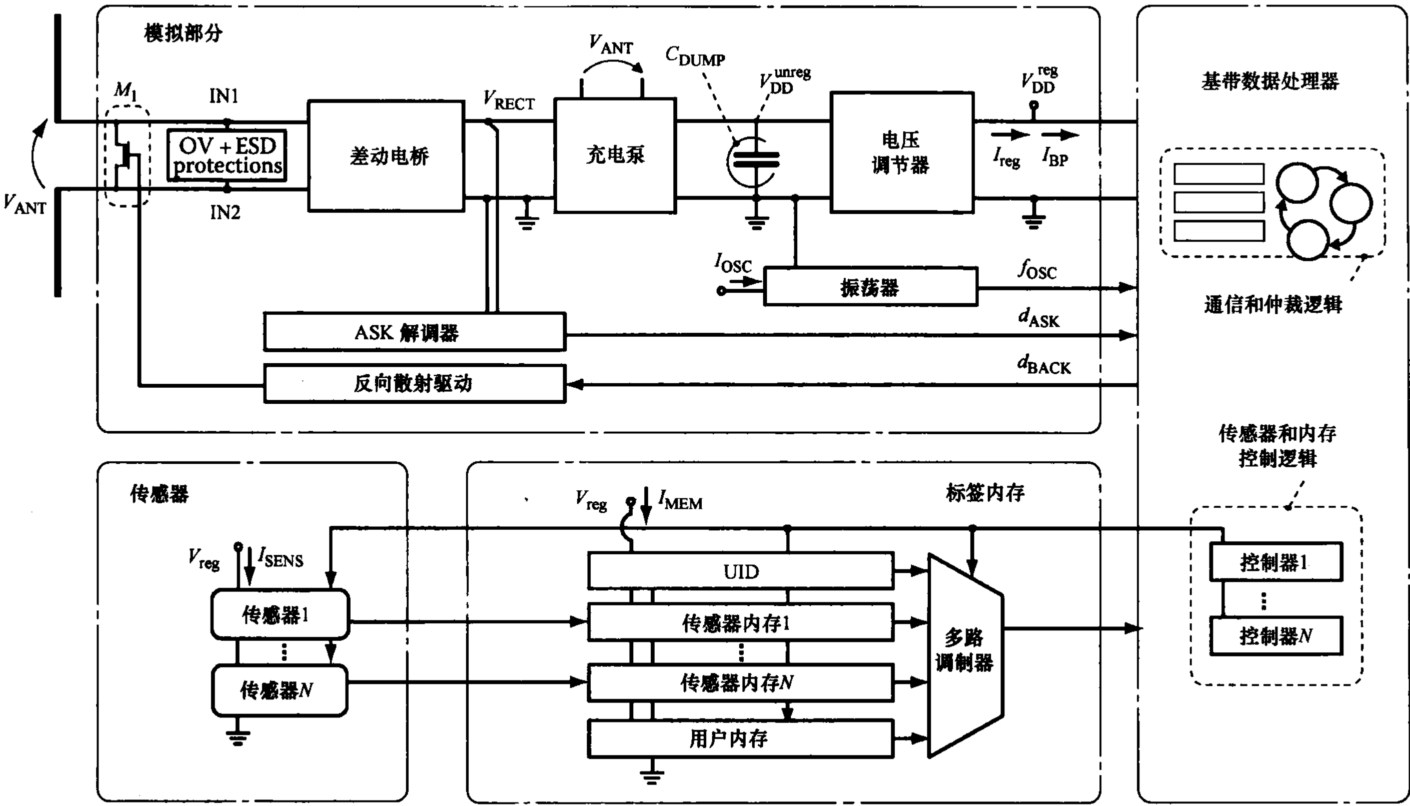


图3-2 UHFRFID转发器框图

被归于标签的不同构件：

$$E_{\mathrm{OUT}}(t)=\int_{t_0}^{t}P_{\mathrm{OUT}}(\tau)\mathrm{d}\tau=\int_{t_0}^{t}[V_{\mathrm{DROP}}(\tau)+V_{\mathrm{DD}}^{\mathrm{reg}}(\tau)]I_{\mathrm{reg}}(\tau)\mathrm{d}\tau \tag{3-9}$$

式中　V_{DROP}——下降到线性调节器的电压。

调节器电流 I_{reg} 等于基带处理器、存储器、传感器和振荡器电流的总和。

$$I_{\mathrm{reg}}=I_{\mathrm{BP}}+I_{\mathrm{MEM}}+I_{\mathrm{SENS}}+I_{\mathrm{OSC.}} \tag{3-10}$$

输入的能量可以表示为

$$E_{\mathrm{IN}}(t)=\int_{t_0}^{t}P_{\mathrm{IN}}(\tau)\mathrm{d}\tau\approx\eta_{\mathrm{RECT}}G_{\mathrm{TAG}}\int_{t_0}^{t}P_{\mathrm{EIRP}}(\tau)\mathrm{d}\tau \tag{3-11}$$

式中　η_{RECT}——校正器的转换效率。

3.3.2　制造过程

正如 3.1 节已经强调的，低成本对于扩大标签的使用来说是非常重要的。在大规模生产中，制造成本往往影响着设计成本，因为根据性能价格比来对实际的开发过程进行评估是非常重要的。就速度而言，RFID 标签通常没有严格的要求，而能量是一个主要的关注因素。进一步的研究将涉及维护等相关技术，因为保证投入成本需要长期的运行技术做前提。另外，采用低成本的、一般的芯片制造技术可以降低每个标签的价格。虽然，模拟部分的设计会受到设备非优化性能和具体处理选择的缺失的影响：这些限制在智能电路级解决方案的设计中必须予以克服。

接下来，当介绍标签原型设计细节时，会提及关于 0.18μm CMOS 技术的节点。这个技术现在已经非常成熟，并能同时确保长期的运行，也不会存在严重的电流泄漏问题。虽然选择的技术可以提供设计者不同特性的阈值电压的晶体管（高阈值电压 - V_{th}=0.5V，低阈值电压 - V_{th}=0.35V）。只有高阈值电压的设备，才在数字核心设计中使用，来最小化对能量消耗有影响的关闭电流的泄漏。调整一个数字核心电压使它接近于晶体管阈值（也就是说，V_{DD}=0.6V），同时限制动态能量消耗，并利用相关的速度需要：仿真和测量数据都支持已选择的方法，甚至在最差的运行状态下。尽管如此，更高技术规格的节点（也就是说，0.13nm）也可以适合物理实施[13]。

3.3.3　空中接口标准

对于 RFID 无源标签的传输，已经有了多个标准。实际上，标签和阅读器在超高频（UHF）频谱下通信的国际标准是 ISO 18000-6 。现在这个标准有三个版本，ISO 18000 -6A，-6B，-6C。本节将主要分析 -6B 和 -6C 标准，它们比 -6A 使用广泛。另外，当标准从一个换成另一个时，芯片的其中一部分设计得到重复使用是有可能的。

-6B 和 -6C 标准都是用来定义无源的反向散射的 RFID 系统，具有以下特性：与区域内的多标签进行认证和通信，选择一个标签子集进行认证或者通信，多次读

取或者写入以及重写单独的标签、用户控制的可以永久锁定的存储器；数据完整性保护；可以检测阅读器到标签链路的错误，以及标签到阅读器链路的错误。

根据ISO18000-6B标准[17]，阅读器调制射频载波信号的振幅（也就是ASK传输）来把数据传给标签。标签接收器可以对经过曼彻斯特编码的数据进行比特错误检测。反向链路空中接口是基于典型的反向散射调制[20]：标签通过调制入射能，并将其反射给阅读器来传输信息给询问者。此标准允许的最大传输速率为40kHz，此速率可以使读标签的速度最大化。协议使用概率性的二进制树算法来处理碰撞。另外，在阅读器和标签之间传递的命令和数据被包含在格式化的帧中，这些帧被分成不同的部分。每个帧包含一个16位的循环冗余校验（CRC）部分，来提高检错能力。图3-3描述了一个兼容ISO 18000-6B的转发器的主要状态。3.5.3节将进一步介绍细节，并讨论传感器的集成。

根据ISO 18000-6C标准，询问器到标签的通信也是基于一个基于包的方案的[23,24]。在ISO 18000-6B基础上，-6C标准支持更高的传输数据的速率：传输速率由询问器定义，并且可以在每次通信循环中更改。为了能够接收脉冲间隔编码（Pulse Interval Encoding，PIE）调制的有效载荷数据（如果有必要的话，还可以响应阅读器），标签需要测量阅读器包头在内的参考时间间隔（Tari，RTcal，如果需要的话，还有TRcal）。Tari间隔的测量（例如0标志的长度，会在{6.25μs、12.5μs、25μs}的这个集合中选择），RTcal（例如数0脉冲或数1脉冲）可以通过至少2MHz的参考时钟频率来计算得到。阅读器可以利用不同的RTcal间隔的组合来指定标签的反向散射链路频率（LF，每隔40÷640kHz的范围内），并分割比例参数（DR，包括有效载荷）。实际LF和它的偏差FT，限制了标签的最小化时钟频率，在3.4.7节将会进行深入的分析。响应识别命令而选择的编码格式，是FM0或是Miller调制子载波。标签的清查循环是使用时隙随机防碰撞算法来进行的。在6C协议中也包括了基于16位或者5位CRC域的检错能力，这取决于阅读器命令。

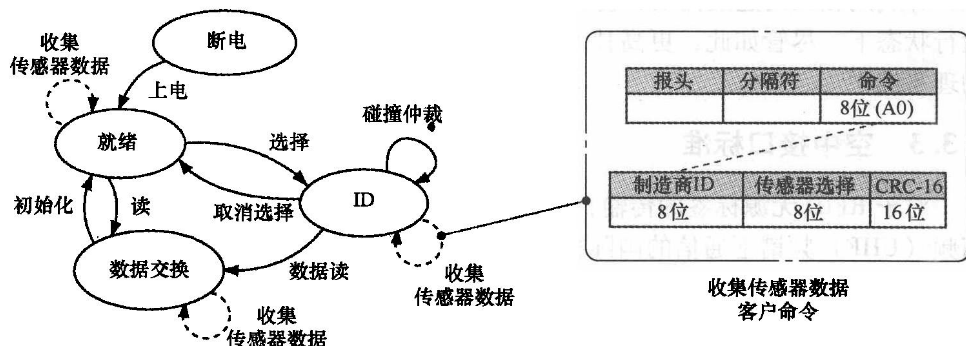


图3-3　ISO 18000-6B标签状态图和用户命令格式

3.4　模拟前端和天线设计讨论

如前面所述，一个无源的 UHF 标签会从阅读器发射的电磁波中获得标签芯片的模拟和数字部分需要的能量。因此，射频前端的灵敏度和能量效率是一个重要的问题，这会直接地影响到标签的识读范围。图 3-2 为标签模拟前端的黑盒原理图。模拟前端通过 IN1 和 IN2 引脚与天线终端连接，IN1 和 IN2 是 RFID 芯片的唯一两个引脚。其中的非平衡桥是一个射频校正器，它用来校正天线的最大能量，并提供一个校正的电压给接下来的模块。由于校正后的电压太低，不能给芯片的数字部分供电（与阅读器相距最大距离时，这个电压只有数百毫伏），所以有必要使用一个电压升压器（参考图 3-2 中的电荷泵）。通过片上滤波电容（ C_{DUMP} ）这个 DC-DC 升压转换器可以产生一个稳定的电压（ V_{DD}^{unreg} ），升压后的电压可以足够天线来使用，提供一个有效的识读距离。因此，需要引入一系列的电压转换器来提供一个几乎独立于识读状态的稳定的电压源，为芯片的数字部分提供能量。限制电压源的变化，可以降低本地振荡器提供的参考频率的偏差，这些将在 3.4.7 节中讨论。

模拟前端的另外一个特性是对阅读器发送的信号进行解调。比如，ISO18000-6B/-6C 规定，阅读器向标签的通信是基于以一个调制深度从 18% ~100% 变化的 UHF 载波的 ASK 调制[35]。最后，模拟前端会以标签向阅读器通信来进行响应，这个通信是基于反向散射技术的[22]。这是一种很好的方法，它可以以最小的能量消耗传输回任意信息：标签会激活或关闭与天线平行的开关（M1）来发送一个 0 或 1 给阅读器。当开关关闭天线时，天线完全不匹配，将会产生从标签到阅读器的一个很强的能量反射。因此，标签位传输的逻辑值可以通过测量反射能量的大小来识别。

3.4.1　天线特性

为了从射频区域获得最大的可用能量，在天线和芯片的射频前端之间必须获得良好的阻抗匹配状态。这样的限制意味着在发射频率处的天线阻抗需要与前端输入阻抗的共轭复数相等[26]。后者由一个电阻性的实部和电容性的负的虚部组成，把芯片等价于一个 $R_{IN}-C_{IN}$ 串联的电路。因此，如图 3-4 所示，可以设置天线的等价电阻 R_{ANT} 与 R_{IN} 相等，C_{IN} 与电感 L 平衡，使能量实现匹配。精确的设计可以使电感的值低到数十纳享，所以不需要放置一个作为外部元件的电感，外部电感将会增加不期望的额外消耗。实际上，可以设置感应天线的谐振频率 f_{ANT} 略高于通信中使用的射频载波（也就是说，f=869.5MHz）：一个设计优良的天线，可以不加任何额外元件，即能形成能量匹配状态。如果实现了匹配状态（即 $R_{ANT}=R_{IN}$ ，并且 $\omega L=(\omega C_{IN})^{-1}$ ），式（3-5）的 Friis 方程可以在不用引入任何由于阻抗失配而引起的衰减因素的情况下使用，并且此时，芯片输入端可用的能量是在给定的距离下最高

的。现在根据图3-4就可以来计算输入峰值电压。公式如下，

$$V_{IN} \approx \sqrt{2\frac{P_{AV}}{R_{IN}}\frac{1}{\omega C_{IN}}} \qquad (3-12)$$

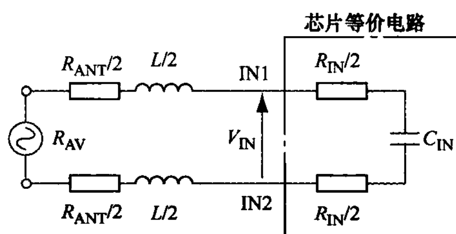


图3-4　被动标签等价图解

从上面的方程中可以看到，设计一个具有低电阻和低电容值的标签，可以提高输入峰值电压值。因此，根据RF/DC能量转换的非线性特征，通过提高输入峰值的水平，可以实现更高的能量转换效率，标签必须通过某种方式，以便最小化R_{IN}和C_{IN}。需要注意的是，一个具有非常低实部阻抗的天线的物理设计，是一个困难的任务，因为此时天线增益将会下降。此外，天线电路中电阻产生的寄生效应，也会导致效率明显地下降。

3.4.2　射频整流器

如上所述，无源标签利用整流和规范射频载波，从电磁场中获得能量。式（3-13）为阅读器发射的没有经过调制的载波：

$$v_{IN}(t) = V_{IN}\sin(2\pi ft) = V_{IN}\sin\left(\frac{2\pi t}{T}\right) \qquad (3-13)$$

式中　f——超高频（UHF）载波频率（比如，在欧洲是869.5MHz）；

T——相应的时间周期。

过去几年，绝大部分RFID标签使用的是全波桥式二极管结构，如图3-5a所示。在RECT端的电压水平为一个二极管的阈值（V_T），低于IN1或IN2能达到的最大潜在峰值。为简便起见，假设负载电容的影响很小。同样的，在GND端设置V_T高于两个输入引脚中的其中一个最小峰值。因此，整流后的电压值V_{RECT}曲线图如图3-6a所示，为

$$V_{RECT} = V_{IN} - 2V_T \qquad (3-14)$$

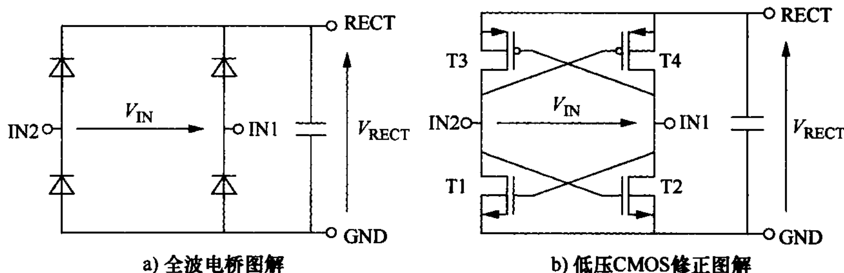


图3-5　全波电桥和低压CMOS修正阶段的图解

从式（3-14）中可以得出，最小输入峰值可以用来实现一个非空整流电压，V_{IN} =

$2V_T$。在整流过程的一个简单步骤中，两个二极管阈值的消耗需要使用低 V_T 的设备，特别是在应用中，V_T 与 V_{IN} 相比是不能被忽略的情况下。事实上，大部分已报道的论文会使用选择具体的技术。大部分已经使用的技术是肖特基设备[5,27-29]，它们的特点是具有较低的正向电压，不过有些参考文献报道了标签采用 BiCOMS[30] 技术，绝缘硅（SOI）[15] 技术，或者是铁电技术。这些可以选择的技术，在设计高能效的整流器时，都具有各自的优点。但是，这些技术的生产成本都高于标准的 CMOS 制作过程。

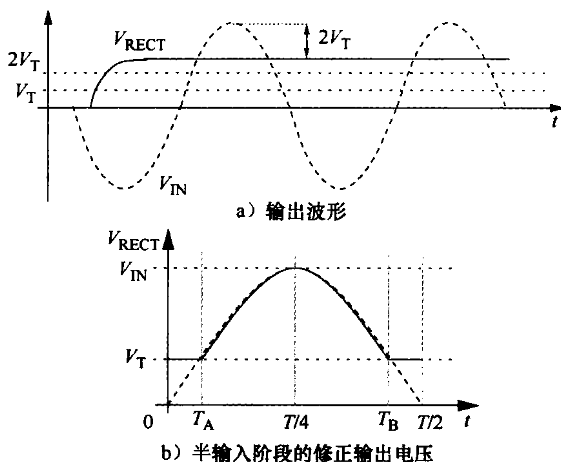


图 3-6 全波电桥

在 CMOS 技术中，由于图 3-5 中的电路的 $2V_T$ 电压降的限制会较大地影响识读的范围。参考文献［25，32，33］报道了一个完全兼容低成本 CMOS 技术的射频整流器。图 3-5b 为相应的电路图。为了能够更好地理解电路的工作原理，首先假设，忽略四个晶体管的电阻，把它们当做是理想的开关器件。此外，认为 PMOS 和 NMOS 的阈值是相等的：$-V_{THp}=V_{THn}=V_T$。这种近似可以使 V_T 是 $-V_{THp}$ 和 V_{THn} 之间最大的。在当 $|v_{IN}(t)|$ 高于 V_T 的时间间隔时，NMOS 晶体管的 T1 或 T2 其中之一的 V_{GS} 电压会高到使晶体管导通。另外一个晶体管，则会有一个值相等的、由电路对称性和输入信号给出的负 V_{GS} 电压，这个晶体管会处于关闭状态。已经关闭的开关会把 GND 与它的源之间进行短路，即引脚具有最低的电压。因此，GND 在输入端具有最小的潜在表现。同样的，RECT 端会与输入引脚相连，表现出或者是 T3 或者 T4 的最高的潜力。当 $v_{IN}(t)$ 是正相期间，关闭的是 T3 和 T4，反向期间，导通的则是 T2 和 T3。即会有，$v_{RECT}(t)=|v_{IN}(t)|$。在应用中，为了使纹波减少到一个可以接受的水平，$v_{RECT}(t)$ 必须要进行滤波：这个过程是在考虑了 MOS 管的电阻 r_{SW} 的情况下完成的，而先前的计算以及被忽略掉。事实上，两个晶体管的电阻 r_{SW} 和负载电容形成了一个低通 RC 滤波器，可以用来减少 UHF 的纹波。在这种情况下，$v_{RECT}(t)\approx V_{RECT}$，后者表示的是整流波形的平均值，它可以根据图 3-6b 计算得到。分析一半的输入周期（例如 $v_{IN}(t)>0$ 的时候）整流器的输出电压可以表示为

$$v_{RECT}(t)=\begin{cases}V_{IN}\sin\left(\frac{2\pi t}{T}\right) & T_A<t<T_B\\ V_T & 0<t<T_A,T_B<t<\frac{T}{2}\end{cases} \tag{3-15}$$

式中

$$T_A = \frac{T}{2\pi}\arcsin(\frac{V_T}{V_{IN}}) \tag{3-16}$$

并且

$$T_B = \frac{T}{2} - \frac{T}{2\pi}\arcsin\left(\frac{V_T}{V_{IN}}\right) \tag{3-17}$$

波形的平均值可以结合式（3-15）计算得到。但是，根据函数的对称性，可以只用这个间隔（即 $T/4$ ）的一半来简化计算，计算结果是

$$V_{RECT} = \frac{2}{\pi}\left[V_T + V_{IN}\cos\ \arcsin(\frac{V_T}{V_{IN}})\right] \tag{3-18}$$

图3-7所示为 V_{RECT} 与 V_{IN}/V_T 比例的关系；在同一个图中，比较了整个波形的直流输出与已整流后波形的关系。假设图3-5a中的二极管和图3-5b中的MOS晶体管的 V_T 是一样的。需要注意的是，采用的方案表现出一个较低的活化阈值（ V_T 对 $2V_T$ ）和一个较高的低输出电压的整流输出，即在 $V_{IN} < 5.65V_T$ 的情况下：这将更适合在最差的运行环境中增强标签的识读范围。事实上，当 $V_{IN} > 5.65V_T$ 时效率会降低（关于全波桥），但不会影响标签的正确行为，因为在这种情况下，由整流器驱动的升压电路可以给芯片提供充足的能量。

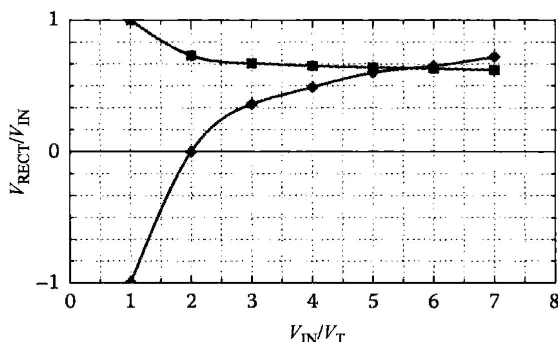


图3-7 全波（◆）与低压CMOS（■）整流器输出的比较

至于嵌入在射频整流器中的 RC 滤波器，它必须降低在 2 × 900 MHz 附近的UHF频率产生的输出纹波，来保护输入RF载波的幅度调制。为了实现这个目的，滤波器的极频率需要位于50kHz ~ 50MHz之间，来实现幅度调制的一个快速路径和对UHF波形的一个较好的抵抗能力。这个频率跨度在一开始时会很宽，元件尺寸的选择将是非常重要的，原因有如下三点。

1）桥式开关的电阻取决于输入电压。实际上，输入电压越高，晶体管的电阻将会越低：考虑在桥中使用的低阈值MOS器件（为了进一步降低激活的最小水平），因为器件上（1.8V）上允许的输入峰值电压在 V_T（0.3V左右）和最大 $|V_{GS}|$ 之间，一个MOS的值 r_{ds} 可以表现为多于二阶的变体。

2）MOS 整流器的小信号等效分析表明，等效的一系列芯片的阻抗会部分地被输入电阻值影响。由于高 R_{IN} 、更高的 R_{SW} 和更高的 R_{IN} 会增强输入数值电压 V_{IN} ，因此有潜力提高效率。此外，电容需要足够大，以便在 UHF 时，有一个较低的阻抗。

3）在先前的计算中，由于电流被估计为几微安，在 MOS 的 r_{ds} 上的电阻损失可以忽略不计。但是，过度地提高晶体管的片上电阻在通过开关时，会造成一个明显的电压降，会掩盖掉先前提到的优点。

3.4.3　电压升压器

先前描述的桥的输出水平常常较低，而不能给电路提供一个有效的电压，特别是当识读距离达到数米的时候。接下来，将介绍另外一个提高直流电压的整流阶段。在可以胜任这项任务的不同架构之间，通常迪克森电荷泵[34]在考虑效率和面积的使用上，可以认为是一个很好的解决方案。为了保持两个输入引脚上负载的对称性，可以考虑使用一个伪结构。图 3-8 所示为一个三阶段的差分迪克森电荷泵。如前面提到的，在低成本技术中使用肖特基二极管是不可行的。为了减少通过二极管时的电压下降，可以使用二极管连接的低阈值 MOS 器件来实现。这个解决方案的效率主要受器件的体效应和基板寄生电容的影响。

图 3-8 所示的三阶段差分迪克森电荷泵，在 IN2 电流低，第一个二极管打开，A1 设置为 $V_{RECT}-V_T$ 的半个输入周期，接下来的半个周期 A1 将与 IN2 一起上升，二极管会关闭。因此，当 IN2 高时，二极管会停止电流流动，IN2 低时又会电流流动。同样的，二极管可以当做为一个开关，当 IN1 低时打开，当 IN1 高时关闭。实际上，根据式（3-18），一般的，V_{RECT} 低于输入峰电压 V_{IN} ，而 IN1 和 IN2 的峰值潜力将高于 RECT 端。如果这些最大电压之间的差高得足以导通一个晶体管，则第一个二极管可以用一个设备来替代，这个设备作为一个由输入最大电压驱动的开关，在不失去阈值电压 V_T 的情况下，可以实现同样的作用。在深亚微米（0.25μm 或是 0.18μm）技术中，本地零阈值的 NMOS 晶体管（ZVT）通常是可行的。这样一个设备，能够成功地来执行任务，允许在整流过程中，获得一个 V_T ：两个 ZVT NMOS 晶体管被放置在第一对二极管的位置，一个由 IN1 驱动（IN1 在 RECT 端和 A1 之间），另一个由 IN2 驱动（IN2 在 RECT 和 B1 之间）。晶体管的大小是由平衡两个限制之后来决定的：①它们必须足够大以忽略掉由于它们自身电阻所造成的电压降，也不会对输入部分增加太多的容抗。②用来提高效率的技术是用一个与 3.4.2 节中桥相似的交叉耦合的 PMOS 峰值检测器来替代最后几个二极管。图 3-9 所示为一个超低电压的电荷泵。

在阅读器向标签和标签向阅读器的通信中，存在一些输入能量突然降低到几乎为零的时间间隔。这发生在标签向阅读器通信中，监测到逻辑 0 的时候，或是复读调制深度接近于 100% 的时候，又或者是在反向散射时，当一个高反射率状态的时

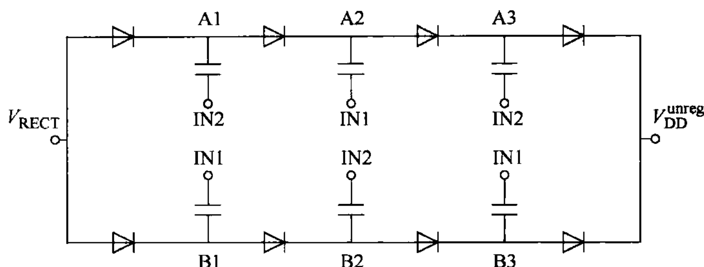


图 3-8　三阶段伪码差分迪克森电压放大器图解

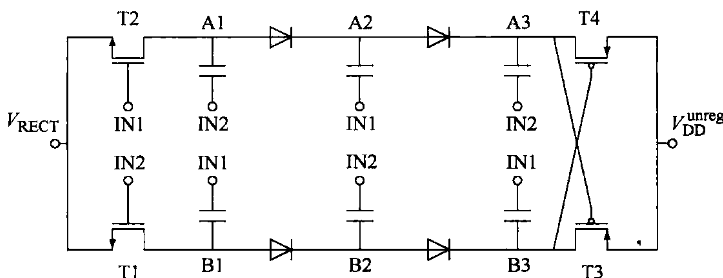


图 3-9　低压放大器图解

候，都可能导致输入能量的不足。正如前面已经强调的，能量存储单元是为了保证从标签激活到通信结束的瞬间，数字核心的能量供给的需要。片上存储电容（C_{DUMP}）的尺寸是在最大芯片面积的基础上，被认为是输入能量不足时的最大时间间隔。此外，可行的标准需要一个标签的最大激活时间（400μs ~ 1ms）：这最终限制了片上电容的最大值。

3.4.4　设备安全保护

有两种电子事件，虽然这两种事件的来源不同，但是都会造成芯片物理上的损坏：一种是静电释放（ESD），另一种由于非常短的读范围而在两个晶体管端部产生过度的潜在差异。当静电事件发生时，如果能在二极管或GC-MOS器件的每对焊盘之间保证低电阻路径，那么ESD的破坏影响是可以避免的。但是，这种方法会在射频输入的焊盘处引入一个较大的电容，会造成输入信号峰值的下降。此外，对图3-5b整流器的分析可以发现，在射频焊盘之间附加一个电容，会造成接收部分输入阻抗的电阻下降，会增加天线设计时的问题。可以造成射频前端永久性损坏的第二个影响是在输入焊盘处的过压，会在标签距离阅读器很近的时候发生。如果深亚微米技术使用的是0.18μm，那么最大的电压是2V。在标签到阅读器的距离大约为1m，阅读器的ERP为500mW时，这种状况将会发生。如图3-10所示，是对具有最小附加电容的RFID模拟前端的适当保护。电路的右边部分，是一个修改后的GC-MOS保护。当一个ESD事件发生时，输入电压 |IN1 - IN2| 将会突然地上升。

用二极管相连的晶体管 T1 ~ T4 会允许 RECT 端和 GND 端迅速地跟着变化。前面上升的部分，将会由 C_C 报告给 T5，T5 导通会突然的打开。在一个 ESD 事件中，会提供 T1-T5-T4 或是 T2-T5-T3 的低电阻路径，这取决于补偿的极性。图 3-10 中电路的左边部分，提供了一个输入过电压的保护。当输入信号的水平到达晶体管 T1、T2 的最大门限值时，T1 和 T2 将会打开，在天线终端引入一个短路，从而限制了输入能量。导致永久性损坏的第三个原因，是在升压器输出端的过电压。这种情况在标签与阅读器相距中等距离，输入的电压水平不足以激活输入保护时发生，但它会造成没有校准的电压，使之超过保护水平。一种可行的措施是基于放置在迪克森电荷泵输出端处的钳位二极管。需要注意的是，在这些节点处附加的电容，在正常工作时不会影响电路的性能的。

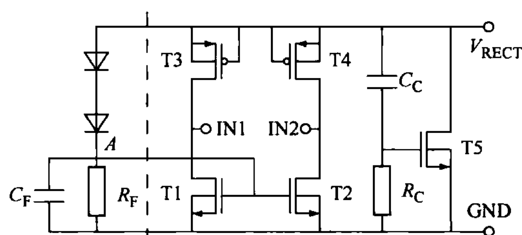


图 3-10 输入保护电路

3.4.5 电压校准

通过能量存储电容的没有经过校准的电压源，取决于可用的输入能量，也取决于识读的距离。为了给数字核心部分提供可以独立于识读条件的稳定电压，需要一个如图 3-11 所示的低能量消耗的输出校准器。一个适合的校准器电路需要具有在数百纳瓦范围的低的能量消耗。一个使用了一个接地门的 ZVT NMOS 用作跟踪电流校准器，因此可以避免需要一个参考基准电流。电压参考模块设置了校准电压的值。参考电路应能在超低能量消耗下承受一个较大的输入电压的变化（0.8 ~2V）。参考文献［35］报道了一个可行的电路方案。

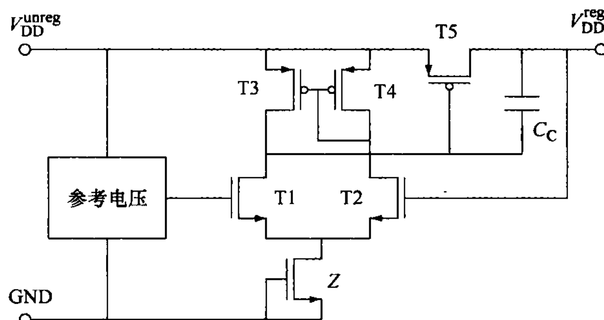


图 3-11 LDO

3.4.6 ASK解调器

3.4.2节介绍了射频整流器，使用小的滤波电容来实现包络检测。因此，在整流器输出端的低频信号是一个阅读器向标签发送的调制信号。由于根据参考标准，调制的深度可以在18%～100%之间变化，所以需要一个电路来将调制信号转换成CMOS数字信号。图3-12所示为一个适合的电路，它利用*RC*滤波器，为比较器的反向输入端提供一个共模级别的振幅调制。

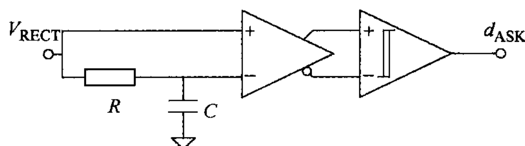


图3-12 CMOS的ASK解调器

3.4.7 时钟发生器

一个兼容ISO 18000-6B和-6C标准的UHF RFID，需要一个本地的振荡器来给芯片的逻辑部分提供一个合适的参考时钟。考虑到低于微瓦的能量消耗、占用小的硅面积和低的输出频率，*RC*振荡器或是环形振荡器对于这样的应用来说，都是合适的结构。已经摆出了影响反向散射信号（δ_f）频率的相关设计，频率必须在实际值的（±15%）/（±22%）之内，这取决于选择的标准和通信数据的速率。这样相对的低值不能用没有调整的*RC*振荡器，或是张弛振荡器来实现。但是芯片种类的调整，将会对芯片的成本有较大的影响，所以应该尽力避免。

在一个环形振荡器中，振荡的频率取决于受MOS管参数（阈值，氧化电容，移动性等）和供电电压影响的转换延时。如图3-11所示，如果振荡器由一个电压校准器供电，由于识读距离改变而产生的供电电压和频率的改变将会被有效地避免。此外，由于影响转换延时的频率变化可以在RFID中引入自动校准函数来进行标记，这可以在数字域内实现。需要注意的是，这种解决方案可以避免任何晶圆分选的附加器件的添加。尽管如此，这种技术要求校整对硅片面积和能量消耗的影响是可以忽略不计的，而且校准不能影响标签的正常运行。最后，校正振荡器需要一个频率参考。

考虑ISO 18000-6B标准的一个时间参考可以从九个连续0的曼彻斯特编码的，数据速率为40kbit/s的识别相位（ASK调制）来实现。本地振荡器可以通过测量由环振荡器提供的频率参考基带信号（由ASK解调器检测）的两个连续上升脚（T_{bit}）的两个时间间隔来进行校准。因此，可以用一个相关容忍度影响的频率参考来测量一个已知的时间间隔T_{bit}。需要注意的是，本地振荡器的频率应当高于$1/T_{bit}$以实现在时间测量中合适的分辨率。这个测量的结果（即在一个时间间隔T_{bit}

内的本地时钟周期数量）随后会用于编程一个放置在本地振荡器和逻辑核心之间的分频器。

需要记住的是，这个被提出的校准方案经过很少的修改，就可以使用在最新的标准中，即 ISO 18000-6C 或者是 EPC1-GEN2。特别是用 PIE 编码的 0 可以在数据的开始处作为时间参考来使用。不过，由最新的标准、范围为 40 ~ 640kHz 的更高的反向散射链路频率来决定的更高的准确度，需要一个更高的晶振频率（几兆赫兹）和 3.3 节介绍的一个更高的编程分频器的分频比率。关于 ISO 18000-6B 的应用，这些修改带来更高的频率和复杂度，因此导致需要更高的功率损耗。在 3.5 节将会详细介绍校准电路（数字时钟管理，DCM）的数字实现。

3.4.8　反向散射发送器

数字部分，一旦设计好，在循环过程中接收到的指示，会提供一个响应的数字信号，并把它发送给阅读器。这个数据流是通过反向散射调制来进行发送的。使用这样的低功耗技术，标签的射频前端会根据发送给阅读器的位逻辑值[22]，在高反射率和低反射率状态之间转换。标准根据参考交叉区域电磁信号的变化，来设置在高和低反射率之间的区别。这个参数依靠天线的物理特性[22]。尽管如此，交叉区域电磁信号较大的变化可以打开通过天线端（即芯片射频前端的输入引脚）的 MOS 开关（可以参考图 3-2中的 T1）来实现。开关的方向性应当最小化，以便于提供一个低的片上电阻，这个片上电阻会导致电磁波交叉区域信号较大的变化。但是，如 3.4.1 节介绍的，输出到基板和源到基板结合处产生的寄生电容会影响射频信号的水平。因此，元件的宽度应当根据最大允许寄生电容的值进行设置。如果开关由水平转换器驱动，并由没有经过校准的电压供电，相对低的片上电阻可以用低的元件宽度来实现。因此，可以用最大化合适的电压来驱动开关的门极。图 3-13 所示为转换器的电路图。

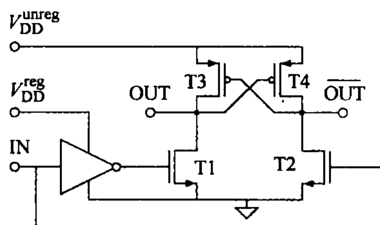


图 3-13　从规则到非规则供应的电平位移器

3.5　数字基带处理器

图 3-2 所示为数字基带处理器的电路结构。数字核心具有多种功能，包括：

1）在输入数据的低频和同步端缩小高速时钟信号的规模。

2）来自阅读器的信息（如曼彻斯特编码数据）的解调。

3）ISO18000-6B 和-6C 分别用两相空间调制或 FM0/密勒调制对反向链路数据进行编码。

4）在前向和反向链路执行帧级别的循环冗余检测。

5）由阅读器命令驱动来进行碰撞仲裁。

6）对传感器设备，包括标签进行管理。

为了实现上述所有的功能，并且使整个能量消耗最小，可以采用多个解决方案，这些方案可以是电路级的，也可以是结构级的。首先，传统的，一般目的的处理器，可以使用 AD Hoc 架构来减少开关活动。另外，基带处理器需要进行优化设计，使驱动的时钟电路保持最低，这样可以减少动态能量消耗。使用大规模、细颗粒的时钟门，可以避免不必要的开关，以便于减少能量浪费。最后，正如已经提及的，使标签核的工作电压略高于阈值电压，在保持平衡性能的同时，能够有效地减少能量的消耗。

已有文献报道了全定制[9]的或是半定制[10-13]的 RFID 基带处理器的应用。前一种方法利用全手动对每一个处理器的建立模块进行优化，来最大化性能（如最小化能量的消耗），但是这样做需要更高的设计代价；后一种设计方式可以显著地减少设计上的成本，但是牺牲了性能上的优化。随着 RFID 技术的演化，接口标准越来越复杂，这些接口标准，使全定制的数字核心的实现变得更加困难。此外，一种基于芯片的设计方法，促进了设计的再利用，这将简化未来节点在处理器方面的设计。为了共轭效率设计和低功耗性能，已经设计了一种标准单元的压缩库，可以适用于设计工作在接近阈值的供电电压的功耗限制的系统。

接下来，将介绍标准单元库的设计。然后，将会描述处理器电路的分区和实现，并考虑低功耗限制和板上传感器集成的相关问题。

3.5.1 低功耗标准单元设计

复杂的数字 VLSI 系统的发展，通常需要依靠一种标准单元设计方式，这种方式可以提供系统物理实现的多个优点。大型的单元库在商业上是可行的，可以在速度和面积的消耗上优化实现高性能。最近，集成电路能耗的减少，成为一个日益重要的问题，制造商开始生产可以用来进行低功耗设计的库。

功耗限制的系统，经常是使用在低成本、低性能的应用中。功耗管理和优化技术是与能量和功耗受限系统不同的。无源标签收发器代表的是功耗受限的系统，在这种系统中，低功耗设计技术需要使用到具体的应用中：这种情况下，长时间的不活动，不会影响功耗预算，而且，如上面的状态，性能（如识读距离）是取决于峰值平均功耗的。

为了说明这些问题的原因，建立一个标准的单元库工作在接近于阈值的供应电压下，目的在于设计一个能量受限的系统。

为了验证提出的方法，需要实现一个受限种类的单元。如参考文献［36-38］所示，使用一个已经选择好的单元的简化集不会影响性能（不管怎样，这不是最主要的），并且可以促进综合处理过程的效率。表3-1总结了一个可行的压缩库的

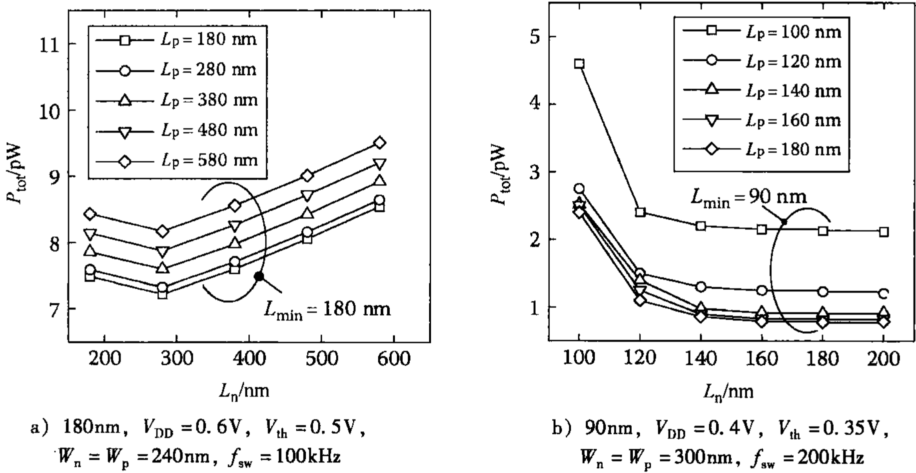
例子，这些例子已经在数个标签实现中得到了应用。双稳态多谐振荡器在被动端与主动端进行转换，使实现部分在时钟信号的每半个周期得到激活，由此，在没有提高主时钟频率的情况下，使运行速度得到了加倍。在整个温度和电压范围内，能够使每个模块得到仔细地设计和展现完全的特性，以便于更好地整合标准库和可靠的数字设计流。

表 3-1　简洁的标准单元库的内容

分类	函　数	单元区域/μm^2
触发器	D-FF 上升沿	68.43
	D-FF 下降沿	68.43
	D-FF 异步复位	131.10
	D-FF 异步调整	131.10
逆变器和缓冲器	INVX1	17.10
	INVX2	22.81
	INVX4	22.81
	BUFX2	22.81
两输入基本	NAND2X1	22.81
	NOR2X1	22.81
	XOR2X1	42.80
多路选择器	MUX2X2	45.86

如上面已经强调的，选择接近阈值电压的方式运行，可以获得超低功耗和平衡的性能。参考 180nm 的实现，单元是在供电电源 V_{DD} = 0.6V 时进行设计的，稍微比高 V_{th} 的元件的阈值电压高，以便于适应电压校准器输出的一些波动。应用于一个 180nm CMOS 处理的高性能系统时，漏电流只扮演一个不太主要的角色，尽管如此，当大部分电路工作在 40kHz 左右的频率时，相应的，漏功耗显得严重了，这将会影响晶体管尺寸的优化。为了调查这个方面，可以通过仿真来找到最优的元件尺寸。图 3-14a 所示为作为一个信道长度的函数的逆变器单元总功耗的仿真。假设开关频率为 100kHz，n-沟道和 p-沟道具有最小化的宽度，使输入电容和动态能量消耗最小。不过，对于一个 180nm 节点功耗函数的中等梯度在位于接近最小化的点附近，这意味着，相对于最小化尺寸元件，选择一个最优化尺寸，预示着仅仅只有 1% 的功耗改善，不能够补偿增加的 18% 的面积。表 3-1 指出了实际的单元尺寸。在这个情况中，最小化尺寸和最小化功耗方案刚好同时出现，这并不是一般的情况。但是，在更低的频率和更低的电路活动状态下，这样的图将会显著地变化，因此需要进行仔细地优化。这个问题随着将来技术的进步，也将会出现。同样的优化过程在设想的 90nm 技术中也进行实施。在这种情况中，假设供电电压为 0.4V，阈值电压为 0.35V，100kHz 的有效开关频率被保留。图 3-14b 描述了这个结果。这个

情况中，最小化的功耗范围与最小化的面积范围相距很远，在分析的范围中，整个功耗可以平均节省70%（虽然增加了面积和速度的支出）。如上面提及的，减少标签的整体功耗，可以提高标签的识读范围，面向应用的功耗优化的重要性变得非常明显。



a）180nm，$V_{DD}=0.6V$，$V_{th}=0.5V$，$W_n=W_p=240nm$，$f_{sw}=100kHz$

b）90nm，$V_{DD}=0.4V$，$V_{th}=0.35V$，$W_n=W_p=300nm$，$f_{sw}=200kHz$

图3-14　一个逆变器单元的功率消耗

3.5.2　基带处理器创建模块

基带逻辑电路设计的目的是保持一个尽可能低的工作频率，并与功能限制兼容。基带处理器部分包括数个时钟域和在时间域的电路活动分布，这些代表了在架构设计层面，使功耗减少的最主要的元素。

最基本的需要是由一些对输入数据的实时处理的需求组成的，也就是说，标签的时钟应当足够快，以提取来自调制射频载波的数据，相关的计算都在空中接口标准允许的时隙中完成。

3.5.2.1　ISO 18000-6B协议实现的方案

参考上面提及的ISO 18000-6B标准，时钟校准电路（DCM）将会在随后介绍，大部分基带电路的工作频率可以低到40kHz。图3-15所示为数字核心的部分结构图。

800kHz的时钟，ϕ_{osc}是相对较快的，它可以由一个小的（有3个阶段的）、高效的环形振荡器有效产生。可以参考DCM，一个小的数字时钟可以用来进行时钟预放大和同步输入的数据（$d_{ask,smp}$），不必使用标签的其他任何功能，可以经受住主时钟频率的大部分波动。

时钟缩减DCM和同步用来获得帧的前端部分，典型的阅读器到标签命令的帧头部，由九个连续的曼彻斯特编码0组成（见图3-16）。

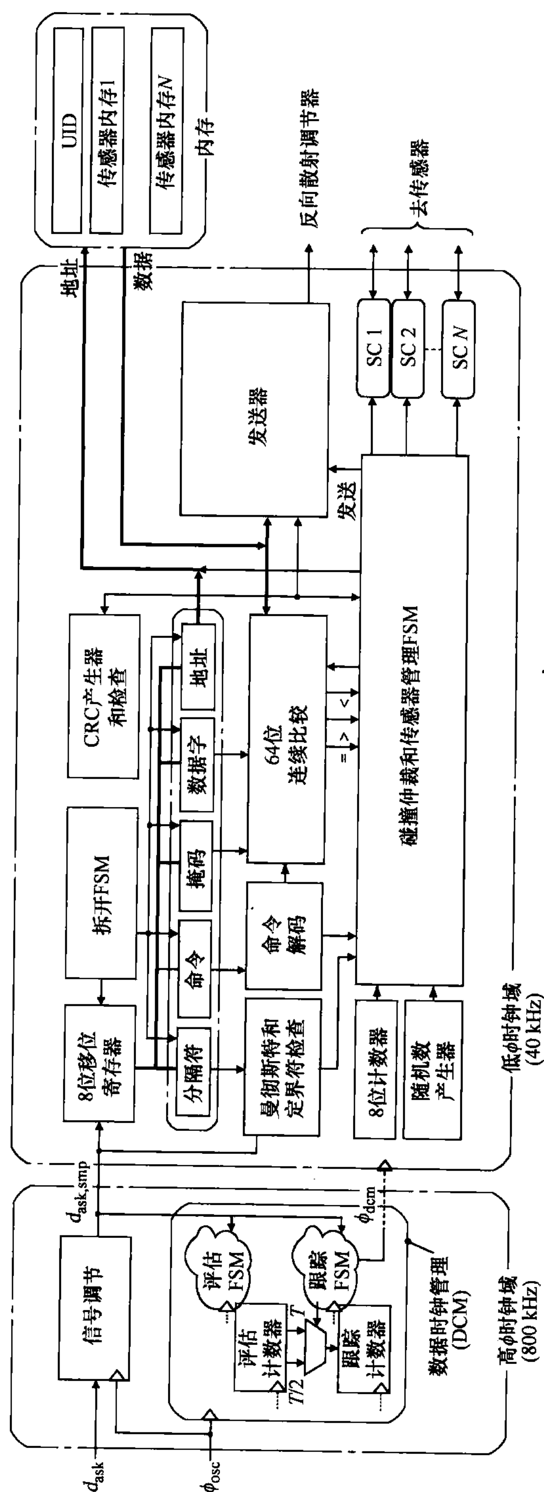


图3-15　基带处理器的结构(ISO 18000-6B协议)

原则上来说，速率为40kHz的曼彻斯特编码的数据流，应当需要80kHz的采样频率。通过使用双边带出发采样，工作的频率可以降低一半，从而减少了能量消耗。反过来，这需要一个输入数据（$d_{ask,smp}$）和预放大时钟信号（ϕ_{dcm}）之间的相近的相位关系，以便保持整个包的分析。这是通过DCM模块来完成的，DCM模块包括两个计数器和一个简单的有限状态机（FSM）。

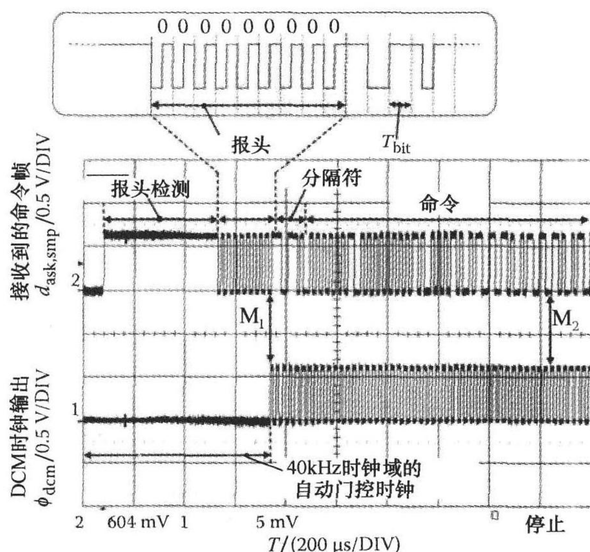


图3-16　基带处理器的结构（ISO 18000-6B协议）

图3-16描述了DCM模块的行为：$d_{ask,smp}$是输入信号，并显示了来自阅读器的一个命令的开始部分。图中可以看到整个部分的前导端和定界符。一旦新数据包的开始部分被识别，间隔时钟（一般振荡频率为800kHz）和输入数据的速率（40kHz）之间的实际比例会通过前导端0标志的平均长度来估计得到。实际上，计算了一个权重的平均，在这之中，权重大的与最后的标志相关：这可以最小化解调器瞬变的影响（这个瞬变在初始位是最敏感的）。在前导端的七个标志位被识别后，将会产生一个脉冲，这个脉冲可以唤醒接下来的电路，并为低速度的时钟信号ϕ_{dcm}提供判决逻辑。从此刻开始，DCM会连续地跟踪输入频率的数据，分析曼彻斯特转换，保证时钟同步，并保持在d_{in-smp}和ϕ_{dcm}之间的一个1/4周期的转换，这对于保证双边沿触发采样的正确性是非常有必要的。位于图3-16波形图中对齐的标记M_1和M_2描述了上述的行为。由一个简单的帧检测器组成的DCM的高频部分可以计算出一个附加的函数，可以用来识别输入命令的开始，并使随后的DCM部分启动。实际上，这允许了由低频时钟驱动的一个自动时钟门控的所有部分。已经用仿真检查了这个方法的可行性和鲁棒性：结果表明DCM模块在甚至更大范围的本地频率时，也可以正确地运行，这也有点不切实际。更精确的，在480kHz和1.320MHz的频率理论范围内获得了同步。图3-17所示为DCM的输出频率（ϕ_{dcm}）依赖于本地振荡器的频率（ϕ_{osc}）。其中，实线描述了仿真行为。

根据

$$\phi_{dcm}=\phi_{osc}/n \tag{3-19}$$

和

$$n = \left[\frac{\phi_{\text{osc}}}{\overline{\phi}_{\text{in-smp}}} \right] \tag{3-20}$$

式中，$\overline{\phi}_{\text{in-smp}}$ 是由输入数据的前导部分计算得到的权重平均值。电路结构的限制 n 的范围在 12 ~32 之间的间隔。低的部分与 DCM 数据跟踪 FSM 相关，这在工作于 $\lfloor n/4 \rfloor > 3$ 是正确的，即 $n_{\min} = 12$，因此 $\phi_{\text{osc}} >$ 480kHz。另一方面，n 的更高部分是与平均计数器的大小相关的，六个位得到保留，用来对两个前导端的位周期进行测量，以限制平均计数器的能量消耗，从而得到 $n_{\max} = 32$，因此 $\phi_{\text{osc}} < 1.320\text{MHz}$。由于门极的延迟和输入数据前导部分的非理想性，测量的性能指出了一个接收输入频率的一个窄的范围，这个范围是 520kHz ~ 1.32MHz。此外，图 3-17 所示为根据 DCM 理论行为测量得到的 DCM 输出频率。ISO 18000-6B 标准定义了反向链路的速率为 40kHz，误差为 ±15%。DCM 准确度满足图 3-17 虚线所示的这个进一步的限制。在此讨论的校准技术中，大部分的数字核心是利用经过校准的低频时钟。振荡器和分频器在正常工作期间运行在高频，因此可以实现较低的能量消耗，正如相关文献所述[39]。

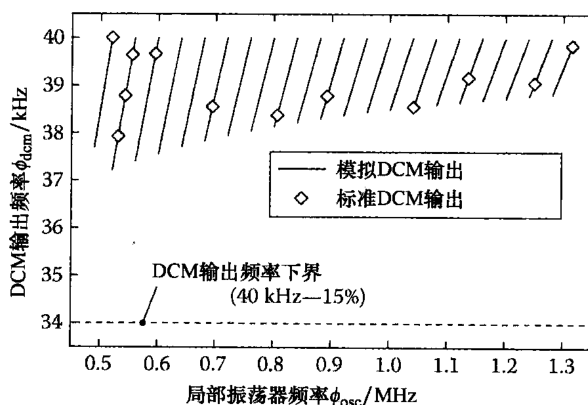


图 3-17 DCM 输出频率：模拟的和标准的行为

一旦输入的数据流被跟踪到以后，根据曼彻斯特和定界符检查模块来计算出数据包的有效性，通过这个检查，可以避免接收一些不完整的数据包。在等待写一个有效的数据包时，时钟将一直保持触发碰撞判断器，来维持 FSM 状态的变化。当一个新的命令帧被识别后，时钟将会短暂的暂停，以便于锁定新的时钟前导。而有效的数据在此时会被转换寄存器处理，进行串并转换，处理后得到八个段。曼彻斯特测试器会连续地监视输入流，帧级别的 CRC 测试用来检查标志的错误。

典型的命令帧包含了多个小的部分和一个较大的部分，包括一个 64 位的字。这取决于实际执行的命令，这个 64 位的字将会以不同的方式与间隔标签码进行比较。

根据能量管理的观点，处理 64 位的字是一个非常关键的问题，这个部分，也会实现一些能耗节能特性。第一点，对输入数据的按位转换将会在 64 位比较器输入端产生一个强烈的开关活动，这是与大电容相关的一个固定特性。为了减小这样的影响，可以考虑使用 8 位的缓冲器，位宽转换也可以实现，从而可以减少由于因子 8 引起的相关功耗。第二点，可以进一步利用相关的弱的时间限制：实际上，64 位操作需要的标准，可以导致一个相当苛刻的组合逻辑网络。尽管如此，可用的时隙允许将 64 位比较操作分裂为 8 个 8 比特的串行化操作，从而对面积和开关电容

（例如能量）进行了一致性的保存。这也使在一些比较模式标准中实现掩藏特性变得简单化：掩藏的字节导致了一个在串行比较中没有操作的循环。

碰撞判决FSM根据ISO的具体标准，管理着整个碰撞判决过程。它需要一个状态计数器和一个随机数产生器。子电路应当独立于控制电路（任何可能的时候，在一个相互专用的方式中），在可能分布的时间内，保持电路的活动。

相同的FSM也同样管理着标签响应的传输。如上面已经提及的，根据两相空间调制，标准会需要反向链路数据以得到编码。根据这样的一个编码方案，一个转换会在每个位期间的结束时发生，并且需要在期间的中部进行一个额外的转换，要求发送一个信号逻辑。因此，在数据速率为40kHz的时候，两相调制标志的转换可能发生于12.5μs的间隔，也就是说，这个频率是有效时钟频率的两倍。为了维持低频率的时钟运行，可以设计一个基于四输入复用器的简单电路方案。图3-18所示为编码电路，它的运行可以很容易地得到解释。平行输入的数据被串行化，然后输入到一个简单的有限状态机中。根据数据位的顺序，传输的信号在连续的高位和低位（“1”数据位）之间的信号间，或是在时钟和反向的时钟信号（“0”数据位）间进行选择。输出的双边沿触发（DET）翻转的目的在于从输出信号中去掉潜在的差错。

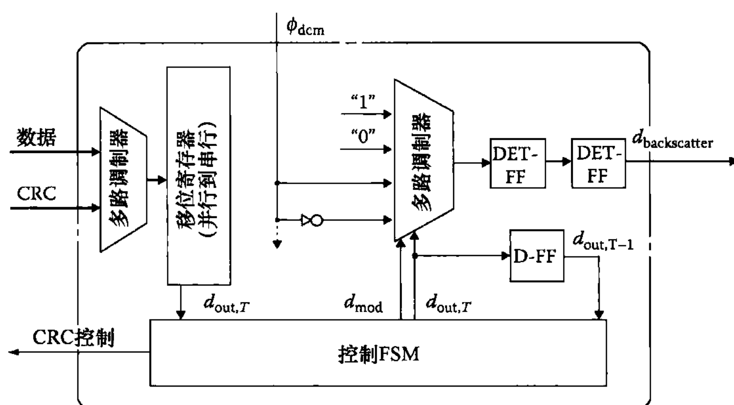
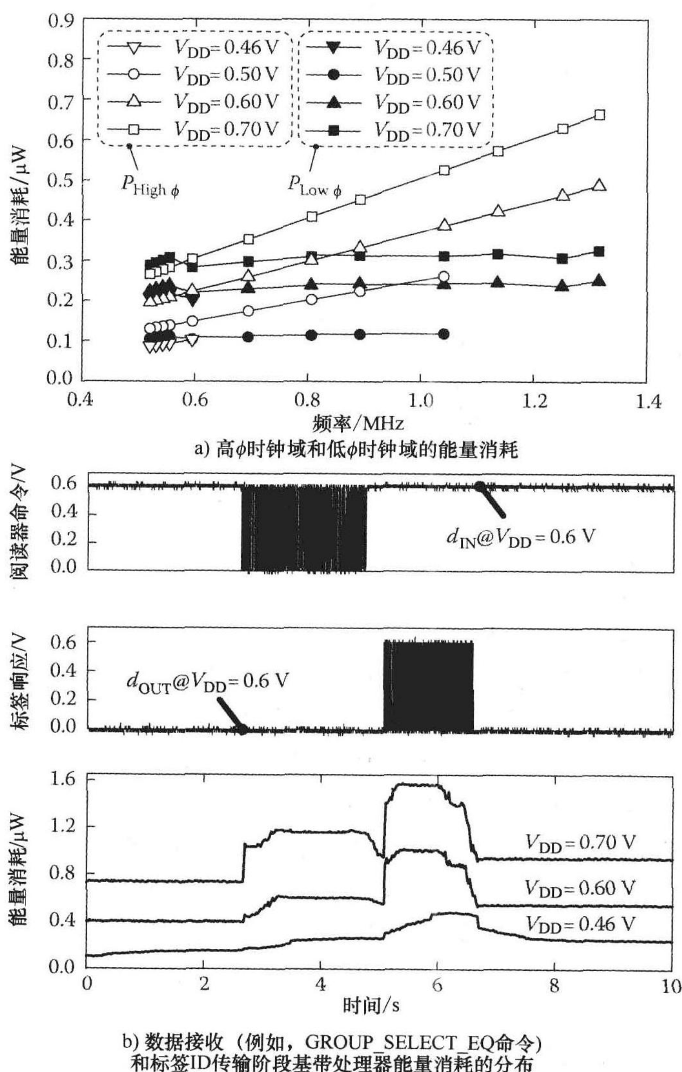


图3-18 发射机组件ISO 18000-6B兼容的细节

最后，数字核的电流状态可以与模拟前端进行通信，来实现整个标签级的节能策略。比如，在时间帧需要来执行标签响应时，ASK解调可以安全地关闭。

正如前面已经提及的，已经制造出了测试芯片，并且进行了一些相关的测试，来验证提出的电路方案。图3-19a报道了在室温下（27℃）测得的基带处理器的能量消耗。实际的供应电流通过不同的供电电压来进行测量，供电电压在{0.46V、0.50V、0.60V和0.70V}这个集合中选择。根据对相对高频部分（比如本地振荡器）和低频部分的贡献，吸收的电流被分成数个部分，以DCM模块作为参考时

钟。正如预期的，高频域能量消耗的增加与时钟速率几乎是呈线性的。当时钟速率超过 600kHz 时，它将高于低频域时的贡献，这个能量消耗将接近时钟频率，正如上面描述的 DCM 模块动作。在目标工作状态下（比如 800kHz 时钟频率和 0.6V 供电电压）数字电路的功耗只有 440nW。图 3-19b 描述了执行循环时能量分布的细节，这与输入信号（例如最上的一个图所示的 GROUP_SELECT_NE 阅读器包）和标签的响应（例如中图所示的 ID 传输）的实际能耗相关。最下面的图显示了根据三个不同供电电压（0.46V、0.60V 和 0.70V）的基带处理器的能量消耗。在响应后进行传输期间，功耗达到了大约 1μW 的最大值。当运行在 V_{DD} = 0.46V 这样低的供电电压时，功耗峰值会只有 400nW。



a) 高ϕ时钟域和低ϕ时钟域的能量消耗

b) 数据接收（例如，GROUP_SELECT_EQ命令）和标签ID传输阶段基带处理器能量消耗的分布

图 3-19　测量结果的例子

3.5.2.2 ISO 18000-6C 实现的方案

正如3.4.7节所介绍的，与ISO18000-6C兼容的时钟规模被标签的回复数据速率限制。利用TR_{cal}间隔和分频比率参数（DR，包含在有效载荷内）的不同组合，识别器具体化了标签的反向散射链路频率（LF，在40kHz ÷ 640kHz的范围内）。实际的LF和它的可以承受的FT限制了标签的最小化时钟频率。参考式（3-21），这样一个频率的估计能够通过仿真来实现。

$$|\mathrm{FT}| = \frac{|T_{\mathrm{LF}}^{\mathrm{nom}} - \hat{n}T_{\mathrm{clk}}|}{\hat{n}T_{\mathrm{clk}}} \tag{3-21}$$

式中，

$$\hat{n} = \begin{cases} 轮\left(\left(3\left|\dfrac{TR_{\mathrm{cal}}}{T_{\mathrm{clk}}/2}\right|\right)\Big/64.2\right) & \mathrm{DR}=64/3 \\ 轮\left(\left|\dfrac{TR_{\mathrm{cal}}}{T_{\mathrm{clk}}/2}\right|\Big/8.2\right) & \mathrm{DR}=8 \end{cases} \tag{3-22}$$

图3-20显示的是DR = 64/3，DR = 8时的结果。反向散射频率（LF）的百分比不能满足所描述的作为本地振荡器的一个函数的频率容忍度的限制。时钟频率应当在1.6～2.20MHz的安全范围内选择。如果一个大致精度的环形振荡器用来作为本地低功耗时钟产生器的话，那么可以采用一个与最小可靠频率相关的最大数值（比如2MHz）。

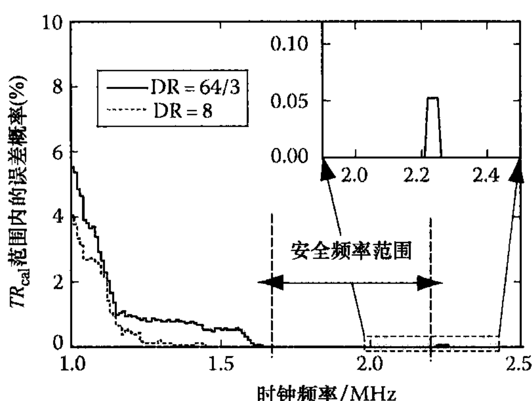


图3-20 ISO 18000-6C基带处理时钟频率选择

图3-21所示为一个ISO 18000-6C基带处理器的可能架构。ϕ_{osc}直接输入给一个PIE解码模块，它可以采样输入阅读器的数据包。解码方式通过数据包界定符认证引发。PIE解码器嵌入了一个二进制的计数器和一些缓冲器，用来测量T_{ari}，RT_{cal}和TR_{cal}。一个简单的FSM使用同样的计数器，来解释连续的阅读器符号，并用$RT_{cal}/2$来比较阅读器的输出。在接收到一个新的符号时，PIE解码器会发布一个T_{OSC}宽脉冲，可以用它来打开其余包的接收电路的时钟信号。当一个6位的计数器（位计数器）计算总的接收符号时，命令解码模块分析第一个已经解码的符号来识别强制命令。10个独立的FSM可以实现管理不同的强制标准命令：使它们保持彼此分离，在稍微增加一些面积的前提下，可以采用更加有效细粒度的时钟门。通过错误处理成分（CRC-16或者是CRC-5，这取决于实际接收的命令），接收的符号流可以用来指出差的PIE编码数据。

标签的状态可以通过一个总是活动的，工作于2MHz的FSM来进行管理。当接

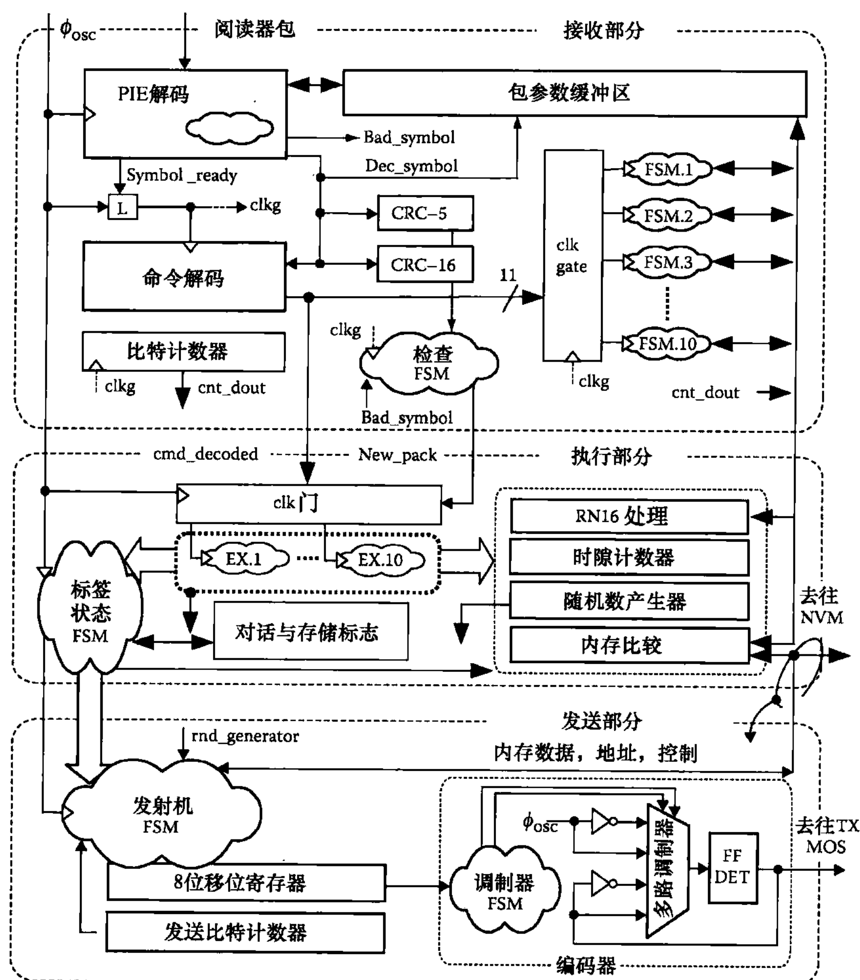


图 3-21　基带处理器的架构（ISO 18000-6C 协议）

收器检查 FSM 有效的输入数据包时，执行会马上开始。其次，不同的命令相关的任务，可以用独立的小控制器来进行管理。执行数据路径包括 4 个独立的模块，用来进行 16 位的随机比较、时隙计数操作、随机产生和存储比较。

在接收数据包和潜在的标签响应之间的可用时间，可以用来把存储比较串行地分成 8 位的操作，因此可以连续地保存面积和开关电容值（比如峰值功率）。

判断过程需要 16 位随机数据，它的产生伴随着一个基于随机产生器的线性反馈转换寄存器（LFSR）。初始化是在上电和标签保持准备时进行的。在判断和发送期间，随机产生器一般会关闭，以便于减少功耗。在判断过程期间，短的再激活（T_{OSC} 的密度）用于进一步地随机数字提取。存储比较在 8 位串行方式中运行，以便于减少面积和开关电容值。

发送器工作包括不同数据（随机数字和存储内容）的FM0和密勒编码。根据这样的方案，一个信号转换发生在每个位周期的末尾；符号“0”（FM0）或者“1”（密勒）是通过在周期中部一个额外的信号边沿来进行编码的。因此，实际的符号转换可能发生在根据选择时钟的双倍频处。

为了避免倍频本地的时钟频率，根据基于四输入复用器（与ISO 180000-6B实现的相似）设计了一个简单的电路方案。图3-21所示为编码器的电路，它的运行也可以直接进行解释：符号边界过渡完成转换输出DET翻转的值，而通过采样简单的脉冲时钟，可以获得中符号转换。一个有限的状态机是否正常驱动复用器，这取决于传送的数据和符号周期。

3.5.3 集成感知设备

兴起的RFID技术旨在片上传感器集成，以形成无线感知设备。与转发器、阅读器和传感外设功耗之间的控制和测量通信相关，传感器的集成有两个主要的考虑因素。ISO 18000-6B空中接口标准支持定制命令的实现，以增加制造商的具体功能。可以定制的唯一部分，是参数和数据部分。任何定制命令应当包括把IC制造商的编码作为它的第一个参数。这可以允许IC制造商在不危害命令码复制的情况下，实现定制命令，因此不会产出误解。这样的命令可以用来管理最终的传感器数据[10]：一个定制码可以把测量的信息保存在标签存储器内，同时，一个标准命令可以应用于采集传感器数据。图3-3根据ISO 18000-6B标准，描述了标签的主要状态：虚线表示了潜在的定制传感器数据采集操作。根据6B标准，阅读器在判断部分前后都可以发布一个定制命令。在前一种情况（“Ready”状态），所有的标签将会涉及潜在的监视任务，而后一种情况（“ID”或是“Data Exchange”状态），只有被选择的转发器参加传感器数据的采集。在图3-3中，定制命令的例子如COLLECT_SENSOR_DATA，可以利用一个传感器选择的8位宽度的区域，单一化或者增加传感器选择。

参考ISO 18000-6C标准，阅读器只有在单独挑出一个标签并知道标签制造商嵌入在标签TID存储器中的识别号之后，才会发布一个定制命令。这限制了潜在监视任务的灵活性，限制了在实际中，对多标签传感数据的同时采集。

从能量管理的观点出发，当标签工作于通信任务（当转发判断，或是接收存储器内容时）时，电源门控可以用来把传感器置于待机状态。通过切断传感器的电流，可以最小化对标签识读范围的影响。

3.6 开放性问题

现在，仍旧很少考虑对于非法RFID标签的清查和跟踪。尽管如此，几年之内，RFID转发器很有可能在多个领域内替代广泛使用的条形码技术，成为供应链

和零售业存储管理的主要部分。因此，提高重要数据的安全性问题希望得到重视。此外，如上面提及的，RFID 转发器可以继承传感器来监视环境和个人的参数。这可能会涉及私人的或是敏感数据的传送，因此会再次考虑隐私问题。为了保证在 RFID 系统内数据的安全性和完整性，必须在 RFID 设备内嵌入合适的特性，来支持数据的隐私和授权[40]。与 RFID 标签的能耗限制是密切相关的加密元效率的设计和实现，能够提供有用的安全工具。

多个工作集中在对标签的隐私保护方案上，这个方案不包括加密元，或者探索专门的 RFID 环境的具体化实现（可以参看参考文献［40］中的一个调查）。最近，欧洲网络在加密方面的优化（ECRYPT）[41]，已经可以识别一个很有前景的新组合密码，专门是针对资源受限的硬件平台，并可能会适用于 RFID 标签。标签上安全特性的实现，以及同时提升性能，仍需要其他研究成果，来进一步扩大 RFID 设备的潜在应用领域。

3.7　结论

本章讨论了无源 RFID 转发器发展采用的技术，特别考虑了直接限制标签性能的能耗问题，讨论了多个能量优化方案，包括模拟前端电路的优化，标准单元的设计和数字系统的能量效率的实现。

讨论了基于 180nm 技术和完全兼容 ISO 18000-6B/-6C 标准的测试芯片的设计和发展。通过采用一个为了数字核心实现的基于标准单元的设计流，可以减少设计的成本，同时为将来的设计提供一个简单的方法和更加优秀的技术节点。同时也讨论了集成传感器的设备和关于空中接口标准的集成。

用仿真和实验测试证实了已经提出的方法的有效性，并得到了运行良好的性能图。通过仔细地选择，能量和面积的平衡可以有效地使用到身边的特殊应用中，通过裁剪设计，使识读的范围最大化。在低成本和高性能的无源标签内，极端的低功耗是可以实现的，以适用于普适计算的应用。

本章工作受到 TECAL 实验室的支持，受到 Regione Emilia-Romagna（意大利）、PRRIITT Misura 3.4 Azione A. 基金的资助。

参 考 文 献

1. K. Finkenzeller, *RFID Handbook, Radio-Frequency Identifications Fundamentals and Applications*, 2nd ed. Wiley, New York, 2003.
2. R. Glidden et al., Design of ultra-low cost UHF RFID tags for supply chain applications, *IEEE Commun. Mag.*, 42(8): 140–151, August 2004.
3. S. Masui, E. Ishii, T. Iwawaki, Y. Sugawara, and K. Sawada, A 13.56-MHz CMOS RF identification transponder integrated circuit with a dedicated CPU, in *Dig. Tech. Papers Solid-State Circuits Conf. (ISSCC)*, pp. 162–163, San Francisco, CA, February 1999.

4. F. Kocer and M.P. Flynn, A long-range RFID IC with on-chip ADC in 0.25 μm CMOS, in *Dig. Papers IEEE Radio Freq. Integr. Circuits (RFID) Symp.*, pp. 361–364, Long Beach, CA, June 2005.
5. U. Karthaus and M. Fischer, Fully integrated passive UHF RFID transponder with 16.7 μW minimum RF input power, *IEEE J. Solid State Circ.*, 38(10): 1602–1608, October 2003.
6. B. Jamali, D.C. Ranasinghe, and P.H. Cole, Analysis of UHF RFID CMOS rectifier structures and input impedance characteristics, in *Proc. SPIE (Microelectronics: design, technology, and packaging II)*, vol. 6035, pp. 313–323, Brisbane, Australia, 2005.
7. T. Umeda et al., A 950-MHz rectifier circuit for sensor network tag with 10-m distance, *IEEE J. Solid State Circ.*, 40(1): 35–41, January 2006.
8. A. Facen and A. Boni, A CMOS analog frontend for a passive UHF RFID tag, in *Proc. Int. Symp. on Low Power Electronic and Design*, pp. 280–285, Tagernsee, Germany, October 2006.
9. V. Pillai et al., An ultra-low-power long range battery/passive RFID tag for UHF and microwave bands with a current consumption of 700 nA at 1.5 V, *IEEE Trans. Circuits Syst.*, 54(7): 1500–1512, July 2007.
10. A. Ricci and I. De Munari, Enabling pervasive sensing with RFID: An ultra low-power digital core for UHF transponders, in *Proc. IEEE Int. Symph. on Circuit and Systems (ISCAS)*, pp. 1589–1592, New Orleans, LA, May 2007.
11. A. Mann et al., Design and implementation of a low-power baseband-system for RFID tag, in *Proc. IEEE Int. Symp. on Circuit and Systems (ISCAS)*, pp. 1585–1588, New Orleans, LA, May 2006.
12. H. Yan, H. Jianyun, L. Qiang, and M. Hao Design of low-power baseband-processor for RFID tag, in *Proc. Int. Symp. on Applications and the Internet Workshops*, Phoenix, AZ, January 2006.
13. R. Barnett, G. Balachandran, S. Lazar, B. Kramer, G. Konnail, S. Rajasekhar, and V. Drobny, A passive UHF RFID transponder for EPC Gen 2 with −14 dBm sensitivity in 0.13 μm CMOS, in *2007 Solid-State Circuits Conference, Digest of Technical Papers*, pp. 582–583, San Francisco, CA, February 2007.
14. W.G. Yeoh, Y.B. Choi, K.Y. Tham, S.X. Diao, and Y.S. Li, A CMOS 2.45-GHz radio frequency identification tag IC with read/write memory, in *Dig. Papers IEEE Radio Freq. Integr. Circuits (RFIC) Symp.*, pp. 365–368, Long Beach, CA, June 2005.
15. J.-P. Curty, N. Joehl, C. Dehollain, and M.J. Declercq, Remotely powered addressable UHF RFID integrated system, *IEEE J. Solid State Circ.*, 40(11): 2193–2202, November 2005.
16. EPC Global, 860 MHz–930 MHz class 0 radio frequency identification tag protocol specification candidate recommendation, Version 1.0.0. MIT Auto-ID Center, June 2003.
17. International Standards Organization, Type B UHF RFID, ISO/IEC WD 18000 Part 6. August 2004.
18. R.J. Marhefka and J.D. Kraus, *Antennas*, McGraw-Hill, New York, 2002.
19. H.T. Friis, A note on simple transmission formula, *Proc. Inst. Radio Eng.*, 34: 254–256, May 1946.

20. ETSI, *TR 101 445*, v1.1.1 ed., April 2002. Electromagnetic compatibility and radio spectrum matters (ERM); short-range devices (SRD) intended for operation in the 862 MHz to 870 MHz band; system reference document for radio frequency identification (RFID) equipment.
21. J. Rabaey, Scaling the power wall, *Keynote presentation, 44th DAC*, June 2007.
22. K.V.S. Rao and P.V. Nikitin, Theory and measurement of backscattering from RFID tags, *IEEE Ant. Propag. Mag.*, 48(6): 212–218, December 2006.
23. EPC Global, EPC radio-frequency identity protocols class1 generation2 UHF, RFID protocol for communications at 860 MHz–960 MHz, Version 1.0.9, January 2005.
24. EPC Global, EPC radio-frequency identity protocols class1 generation2 UHF, RFID conformance requirements, Version 1.0.2, January 2005.
25. A. Facen and A. Boni, CMOS power retriever for UHF RFID tags, *IET Electron. Lett.*, 43(25): 1424–1425, December 2007.
26. K.V.S. Rao, P.V. Nikitin, and S.F. Lam, Impedance matching concepts in RFID transponder design, in *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 39–42, Buffalo, NY, 2005.
27. K. Seemann, F. Cilek, G. Hofer, and R. Weigel, Single-ended ultra-low-power multistage rectifiers for passive RFID tags at UHF and microwave frequencies, in *Radio and Wireless Symposium*, pp. 479–482, San Diego, CA, January 2006.
28. N. Tran, B. Lee, and J. Lee, Development of long-range UHF-band RFID tag chip using schottky diodes in standard CMOS technology, in *Radio Frequency Integrated Circuits (RFIC) Symposium*, pp. 281–284, Honolulu, HI, June 2007.
29. A. Navarro and J.L. Del Valle, Voltage generator for UHF RFID passive tags using Schottky diodes based on a 0.5 μm CMOS technology, in *3rd International Conference on Electrical and Electronics Engineering*, pp. 1–4, Veracruz, Mexico, November 2006.
30. G. De Vita and G. Iannaccone, Ultra low power RF section of a passive microwave RFID transponder in 0.35 μm BiCMOS, in *International Symposium on Circuit and Systems*, vol. 5, pp. 5075–5078, Kobe, Japan, May 2005.
31. H. Nakamoto, D. Yamazaki, T. Yamamoto, H. Kurata, S. Yamada, K. Mukaida, T. Ninomiya, T. Ohkawa, S. Masui, and K. Gotoh, A passive UHF RF identification CMOS tag IC using ferroelectric RAM in 0.35 μm technology, *IEEE J. Solid State Circ.*, 42(1): 101–110, January 2007.
32. A. Facen and A. Boni, Power supply generation in CMOS passive UHF RFID tags, in *PhD Research in Microelectronics and Electronics*, pp. 33–36, 2006.
33. S. Mandal and R. Sarpeshkar, Low-power CMOS rectifier design for RFID applications, *IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications*, 54(6): 1177–1188, June 2007.
34. J.F. Dickson, On-chip high-voltage generation in NMOS integrated circuits using an improved voltage multiplier technique, *IEEE J. Solid State Circ.*, 11(3): 374–378, June 1976.
35. G. De Vita and G. Iannaccone, A sub-1V 10 ppm/C, nanopower voltage reference generator, *IEEE J. Solid State Circ.*, 42(7): 1536–1542, July 2007.
36. N.M. Duc and T. Sakurai, Compact yet high-performance (CyHP) library for short time-to-market with new technologies, in *Proc. 5th Asia and South Pacific Design Automation Conf. (ASP-DAC)*, pp. 475–480, Yokohama, Japan, January 2000.

37. J.M. Masgonty, S. Cserveny, C. Arm, P.D. Pfister, and C. Piguet, Low-power low-voltage standard cell libraries with a limited number of cells, in *Proc. Int. Workshop—Power And Timing Modeling, Optimization and Simulation (PATMOS)*, Yverdon-Les-Bains, Switzerland, September 2001.
38. A. Ricci, I. De Munari, and P. Ciampolini, An evolutionary approach for standard-cell library reduction, in *Proc. 17th ACM Great Lakes Symp. on VLSI*, pp. 305–310, Stresa (VB), Italy, March 2007.
39. F. Cilek, K. Seemann, G. Holweg, and R. Weigel, Impact of the local oscillator on baseband processing in RFID transponder, in *International Symposium on Signal, Systems and Electronics*, pp. 231–234, Montréal (Québec), Canada, 2007.
40. A. Juels, RFID security and privacy: A research survey, *IEEE J. Select. Areas Commun.*, 24(2): 381–394, February 2006.
41. Stream Cipher Project Web Page, ECRYPT (European network for excellence in cryptology), 2005, available: http://www.ecrypt.eu.org/stream/(online).

第 4 章　RFID 的 EPC Gen-2 标准

无线射频识别（RFID）系统存在众多的通信协议。其中一个最流行的协议就是 EPC Gen-2 协议。这个协议在零售业中得到了广泛的使用。Gen-2 协议在物理层特性和链路层过程中提供了较大的灵活性，以适应不同的环境。这个灵活性对于最大化吞吐量或每秒接入/读标签的数量来说，是非常重要的。本章将描述 Gen-2 协议的物理层和数据链路层的概貌，并对最大化吞吐量的不同方案进行探讨。

4.1　概述

无线射频识别由两种设备组成：RFID 阅读器（有时也可称为识别器）和 RFID 标签。RFID 标签附着在一个物体上，并且包含着这个特定物体的信息。典型地，在阅读器、中心数据库和后台系统之间有一些软件应用。本章将集中关注在标签和阅读器通信中使用的一种通信协议。RFID 系统分为两类：标签上没有任何的板上电池的被动系统和标签上有板上电池的主动系统。根据阅读器和标签之间通信的频率，可以对上述提到的两类进行进一步的细分。在 RFID 标签和阅读器通信链路可以使用许多不同的标准，本章将集中叙述 EPC Gen-2 标准。

RFID 与传统的条形码相比，存在着明显的优势。现今，条形码被用来识别物体，但条形码阅读器和条形码在可视范围内才能通信。而 RFID 标签在其可视范围之外也可以通信。此外，RFID 标签可以存储比传统条形码更多的信息，能够用来识别唯一的物品。比如，货号为 1275 的蓝色的牛仔裤。而且，最重要的是，RFID 标签在其使用寿命内，可以写入新的数据，或者更新旧数据。由于这种能力，在 ePedigree 应用中，RFID 标签具有许多的优势（ePedigree 是一种监护财产链的电子日志，例如制药产业）。

4.1.1　EPC Gen-2 背景

EPC Gen-2（Gen-2）标准是 RFID 系统中使用的标准之一。最近，Gen-2 系统被批准为 ISO 标准，ISO 18000 的 6C 部分是被动（没有电池的标签）超高频（UHF）RFID 系统的一个主导标准。许多主要的零售商，如沃尔玛、麦德龙等使用基于 Gen-2 标准的 RFID 系统。

RFID 阅读器和标签在超高频范围内，遵从 Gen-2 标准进行通信，超高频的频带在 860 ~ 960MHz 内，这取决于相应的地理位置。例如，欧盟允许的频率范围为 868 ~ 870MHz，美国允许的范围为 902 ~ 928MHz，而日本允许的范围为 902 ~

960MHz。Gen-2标准定义了阅读器和标签之间的通信，不用考虑系统工作于特定的频率范围

4.1.1.1 Gen-2标准的目标和需求

Gen-2标准的首要目标是提供一个统一的方法来读取来自标签的数据、向标签写数据和与标签通信。目标是希望能够防止电子商品防盗系统（EAS）商场中存在的一些问题。在EAS市场，没有一个单一的工作标准要求制造商和零售商来购买和维护几个不同的EAS设备集和多个不同的EAS标签集。这将导致一个较大的成本。例如，Gen-2这样的标准，通过提供给开发者一张基本的描述标签和阅读器之间通信的基本设计图，描述系统（阅读器和标签之间）如何交互，来解决上述困难。开发者随后能够按照标准来生产产品，这个标准是和其他生产商按照同一标准生产的产品共同使用的，并可以与其他生产商生产的基于统一标准的产品进行通信。因此，使用者可以购买单一的系统，而这个系统可以在全部生产商的产品上工作。

RFID标准面临着众多的问题。第一，必须决定设备工作的频率范围。这个问题是非常难解决的，因为使用的射频是当地专业行业协会控制的。因此，一个RFID标准必须识别世界范围内可以免费使用的频率范围。Gen-2使用一个100MHz的宽范围，而不是一个单一的频带，这是考虑到上述提及的不同的地方规定。第二，通信协议必须满足一般的应用，同时在应用空间内提供专业化的操作。通信协议必须足够简单，以便于提供基本的功能，访问和读写操作。标准也必须提供开发者使用的接口，以便于在基本功能之上提供客户定制的特性。第三，RFID标准必须为每一个标签提供一个唯一的ID号，并且提供一种方法允许这个ID号扩展成更大的数字（每个标签有一个ID号）。第四，当在选择工作频率范围时，也需要考虑系统的物理性能。识读范围便是其中一个物理参数，库存管理系统中，便需要一个长的识读范围，而在传输相关的选择系统中，则需要短的识读范围。

在零售业使用的RFID系统中，Gen-2是一个主导标准。零售业需要一个作为通行证的类似牌照的标签，这个标签可以提供一个唯一的ID号来作为密码访问包含例如成本或是过期日之类信息的中心数据库。Gen-2提供了一个健壮的机制，EPC序号，这个唯一的ID号可以满足零售商需要数以百万计的唯一序列号的要求。此外，一个范围集内的EPC序号，允许使用者快速地识别分配给具体产品的标签。在其他应用空间内，有一些其他的RFID标准，例如图书馆出入管理系统，公交收费系统，或是高速公路和桥梁的收费系统。

4.1.1.2 EPC编码系统的目标和需要

EPC编码系统的首要目标是提供一个统一的系统，在这个系统中，一个特定的标签可以唯一地被识别到。EPC序号包含了识别制造商、物品类型和一个在此制造商和物品类型下的唯一的序列号。因此，EPC序号是普遍的，并且可以基于多个或者仅一个序列号来对标签进行寻址。使用EPC序号后，一个特定制造商的所有标

签都能够被识别，简化了库存管理。EPC序号编码的信息可以用来访问更大数据库内特定物品的信息。

4.1.2 Gen-2通常使用的特性的概述

Gen-2协议使用最普遍的特性是清单命令。库存命令由以下四个命令组成：选择、查询、查询重复和查询调节。这些命令用来读取在阅读器范围内所有标签的EPC序号。EPC序号可以用来访问中心数据库，获得例如标签附着在哪一个物品上等更多信息。在某个环境中，Gen-2标签用来管理清单命令，并可以让使用者在特定的时间内获得手边所有物品的清单。在零售业环境中，这样的信息对于满足零售商的需求是非常有效的。Gen-2协议的第二个普遍的特性是读和写命令。这两个命令是属于访问命令。写命令能够使用户在标签上指定的位置写一个字（16位）。同样的，读命令能够使用户能够读取标签256字的数据。尽管如此，存储器内可能没有包含256字的数据。

并不是所有的零售设施都可以使用RFID，于是RFID Gen-2标签常常是附着在打印的标签上。在打印过程中，RFID标签通过一个打印机运行，这个打印机可以打印一个条形码和标签部分的其他信息。这样可以让不能使用RFID的设备使用条形码识别物品，并可以在RFID标签失效时提供备份。在打印的过程中使用写命令，来向标签写入一个合适的EPC序号。在打印过程中，标签需与一个具体的物品相关联，关于这个具体物品的EPC序号将会写入此标签。EPC序号也可以识别物品的制造商。打印机使用读命令来读回EPC序号，并可以验证写入数据的正确性。EPC序号中附加的其他数据，比如食品过期日，可以使用打印机写入和验证。

4.2 物理层通信特性

Gen-2的物理通信接口与七层开放系统互联（OSI）模型的物理层的概念相似。阅读器控制Gen-2协议物理层的所有部分，并编码发给标签的所有命令的前端部分。在Gen-2协议中，存在着两个通信链路：阅读器向标签的链路和标签向阅读器的链路。这两个链路是互相独立的，存在着不同的数据编码、数据速率和数据调制方案。这两个通信链路的具体特性都由阅读器控制。

这可以使阅读器根据环境的改变调节通信链路。例如，阅读器在射频噪声较大的环境中，可以使用密勒编码来减少标签响应时错误位的数量。又或者是使用者想快速获得大量标签的信息时，阅读器可以使用阅读器向标签链路和标签向阅读器链路允许的最快的速率。数据速率的观点将会在4.2.1节讲述，在4.2.2节讲述不同的调制类型，4.2.3节描述FM0和密勒编码。

4.2.1 数据速率

Gen-2 协议定义了两个通信链路。第一个链路是阅读器向标签的链路，它被用来从阅读器向标签发送命令。阅读器发送命令，随后保持一个载波（CW）。这个载波是没有调制的信号，简化了向标签传输的能量。第二个链路是标签向阅读器的链路，它被用来发送从标签回复的数据给阅读器。在 Gen-2 中，标签与阅读器通信使用反向散射的形式。

当暴露在射频环境中时，所有的天线会吸收环境中的部分能量，并反射剩余的部分。反向散射定义为从任意天线反射的能量。标签具有在两种设置的天线特性之间转换的能力：①反射非常少的能量；②反射几乎全部的能量。当响应一个阅读器时，Gen-2 标签使用这个天线转换能力，调制阅读器发送的载波，编码载波中响应的数据。

4.2.2 调制类型

调制定义为如何把数据在物理层面上编码到载波信号上。Gen-2 支持两种类型的调制：振幅键控（ASK）调制和移相键控（PSK）调制。

如前面提及的，在 Gen-2 中存在着两个独立的通信链路：阅读器向标签的链路和标签向阅读器的链路。阅读器向标签链路的调制，可以使用 ASK 调制的三种类型中的一种：①单边带幅度键控（SSB-ASK）调制、②双边带幅度键控（DSB-ASK）调制或者是③相位倒置幅度键控（PR-ASK）调制[1]。

在每个命令之后，阅读器会发送没有调制的载波信号。标签随后会如 4.2.1 节描述，使用 ASK 或 PSK 来调制反向散射信号。

当一个阅读器只对单个标签通信时，Gen-2 协议是一个半双工协议。因此，当标签通过反向散射发回它的响应时，阅读器不会发送命令。ASK 调制改变正弦波的幅度来区分高低符号。例如，一个正弦波定义如下

$$(m_{\mathrm{d}})(A)\sin(2\pi ft) \tag{4-1}$$

式中 m_{d}——调制深度；

A——正弦波的最大幅度；

f——没有调制的载波的频率；

t——在时间上的分散点。

调制深度 m_{d} 是最大幅度 A 的百分数，由正弦波产生，并且 m_{d} 的范围在 0~1 之间，或者说是在最大幅度与最小幅度之间。当 m_{d} 等于 1 时，正弦波的幅度在它的最大处，这意味着是一个高标志位。当 m_{d} 小于 1 时，正弦波和传输的信号被衰减，幅度也将小于 A，当 m_{d} 低于一个阈值，典型值为 0~0.2 之间时，这就代表着一个低的标志位，这类标签 m_{d} 的最大值会在高和低标志位之间区分，这取决于接收器的灵敏度和标签接收器的电路。

PSK 调制改变正弦波的相位来编码 0 和 1。

4.2.3　数据编码

阅读器发送给标签的数据（阅读器向标签的链路）是使用脉冲间隔编码（PIE）来进行编码的[1]。PIE 使用两个不同长度的脉冲来代表一个逻辑或数据 1 和一个逻辑或数据 0。在 Gen-2 中，数据 0 的脉冲比数据 1 的脉冲要短[1]。在 Gen-2 编码中，*Tari* 是一个基本的参考时间单元，在 6.25 ~ 25μs 之间波动[1]。数据 0 符号在长度上是 1*Tari*，如图 4-1 所示，由跟随在一个衰减载波（CW）之后的发送载波（CW）组成。

数据 0 编码的衰减载波（CW）部分的长度用脉宽（PW）参数来定义。脉宽参数以 μs 进行排序，大小如下式所示[1]：

$$\max(0.265 \times Tari, 2) \leqslant PW \leqslant 0.525 \times Tari \tag{4-2}$$

数据 1 标志位比数据 0 标志位长。具体地，数据 1 标志位的长度是数据 0 长度的 1.5 ~2 倍之间[1]。数据 1 标志位如图 4-2 所示。



图 4-1　数据 0 的 PIE 编码

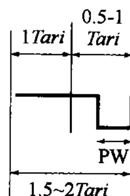


图 4-2　数据 1 的 PIE 编码

4.2.4　信息报头

在无线通信中，一个报头通常标志着一个无线信息的开始。报头具有多个目的。首先，它通知接收器信息正在被发送。第二，它可以使接收器电路自身得到同步，或者锁定信号。第三，报头包含阅读器或是标签这样的信息源的信息，或者是信息编码本身相关的信息。

Gen-2 协议定义了两种报头的集合。第一种报头的集合处理先于阅读器向标签的命令。阅读器向标签命令的报头的类型，取决于发送的命令。阅读器向标签两种类型的报头包含了标签如何回复命令的信息。第二种报头的集合处理标签向阅读器的响应。标签向阅读器响应的这种报头类型取决于响应中数据的编码。

4.2.4.1　阅读器向标签的报头

阅读器向标签的报头包含了标签如何反向散射响应给阅读器的信息。阅读器到标签具有两种类型的报头：报头和帧同步。发送的命令定义使用的报头的类型。当

阅读器发送查询命令[1]时，将会使用报头。除了查询命令之外的其他命令，将会使用帧同步。

帧同步由三部分组成：①分隔符、②数据 0 位和③阅读器向标签的校准（RT-cal）标志位[1]。图 4-3 所示为阅读器向标签的帧同步。

分隔符是帧同步的开端，并具有固定的长度[1]。一个数据 0 位跟在分隔符之后[1]。RTcal 标志位跟在数据 0 位之后，是帧同步最后的标志位[1]。RTcal 标志位的长度与数据 0 位的长度加上数据 1 位的长度相同。RTcal 标志位的长度在 Tari 标志位单元中测量[1]。

标签用 RTcal 标志来区分数据 0 位和数据 1 位。标签用 RTcal 标志来计算支点值。支点值被定义为[1]

$$\text{pivot} = \frac{\text{RTcal}}{2} \tag{4-3}$$

标签接收的比支点短的任何数据位将会被解释成数据 0；接收的数据比支点长的，将会被解释为数据 1[1]。如图 4-1 和图 4-2 所示，数据 0 和数据 1 具有不同的模式。标志也使用 RTcal 来区分有效数据位和无效数据位，或者是正确和错误的数据。标签接收的大于 4 倍 RTcal 长度的任何符号位，将会被解释为一个无效的、错误的数据位[1]。

帧同步过程处理除了查询命令之外的其他全部命令。当阅读器发送查询命令，将会使用报头，而不是使用帧同步。报头包含了 4 个部分：①固定长度的分隔符，②一个数据 0 位；③RTcal 符号和④标签向阅读器的校准（TRcal）符号[1]。因此，报头由紧跟在 TRcal 符号之后的一个帧同步组成。图 4-4 所示为阅读器向标签的报头。

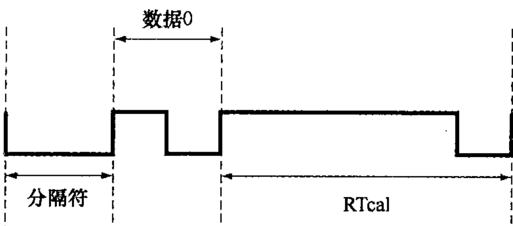


图 4-3　阅读器向标签的帧同步

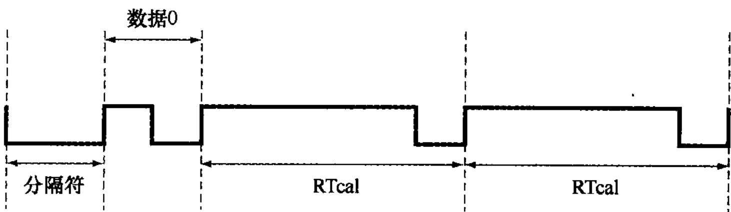


图 4-4　阅读器向标签的报头

分隔符、数据 0 位和 RTcal 符号在帧同步中是区别使用的。在帧同步相同的方式中计算和使用支点值。额外的部件 TRcal 用来指导标签使用什么编码数据和使用什么数据速率。还记得标签向阅读器通信的链路可以使用 FM0 或者是密勒编码么，这两种编码都可以工作于不同的数据速率上。

报头只在阅读器传输一个查询命令给标签时使用。查询命令有一个分隔比例（DR）区域。查询命令中的 DR 区域和前段的 TRcal 符号用来计算标签向阅读器链路的数据编码速率。标签测量 TRcal 的长度，并使用在查询命令中 DR 区域的值来计算数据编码速率，称为反向散射链路频率（BLF）。BLF 由式（4-4）计算得到，

$$\mathrm{BLF} = \frac{\mathrm{DR}}{\mathrm{TRcal}} \tag{4-4}$$

使用 BLF，标签可以知道响应阅读器所需要的编码速率。BLF 在查询阅读器向标签报头的每次查询回合中被设置一次[1]。改变 BLF 的话，需要将 BLF 变为一个新的查询命令，因此需要开始一个新的查询回合。

4.2.4.2　标签向阅读器的报头

标签通过一个报头先于所有的响应。标签向阅读器链路的数据编码可以选择 FM0 和密勒编码之中的一种。使用的报头取决于在响应中数据编码的类型。阅读器通过选择查询命令[1]的 M 区域来控制使用数据编码的方案。

标签报头的导频音可有可无。当标签响应一个命令，并把它写入存储器时，导频音总是包含在标签的报头；否则，支点的包含取决于查询命令 TRext 区域的值[1]。当查询命令 TRext 区域的值为 1 时，将会包括导频音；当 TRext 区域的值为 0 时，将不包括导频音[1]。

FM0 报头由紧跟 FM0 报头的可选的 FM0 导频音组成。FM0 导频音由 12 个 FM0 编码的 0 组成[1]。FM0 报头包含一个设置的数据序列和一个违反 FM0 编码的位[1]。FM0 报头由六个符号组成。前面四个符号由交替的序列 1 和 0 组成，第一位是 1，紧跟这四位后的是一违反 FM0 编码的位，最后是一个一位来结束报头。违反 FM0 编码的位是一个在位时间中间不需要相位反转的数据位 0，是 FM0 报头第二位到倒数第二位[1]。

图 4-5 所示为没有导频音的 FM0 报头，违反 FM0 编码的位以“m”标记。图 4-6所示为具有导频音的 FM0 报头；违反 FM0 编码的为也以“m”标记。

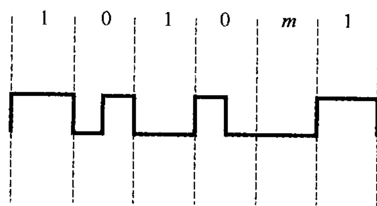


图 4-5　没有导频音的 FM0 报头

密勒编码报头可能被放置在基于 TRext 区域值和阅读器发布的命令的一个导频音之前。密勒导频音由 12 个基带密勒编码的 0 组成，已经与适当的密勒子载波时钟混合[1]。这与 FM0 导频音相似，除了位 0 是根据密勒方案，而不是 FM0 方案进行编码以外。报头有 10 位的长度。不像 FM0 报头，在密勒报头中的 10 位，没有一位是违反密勒编码的。

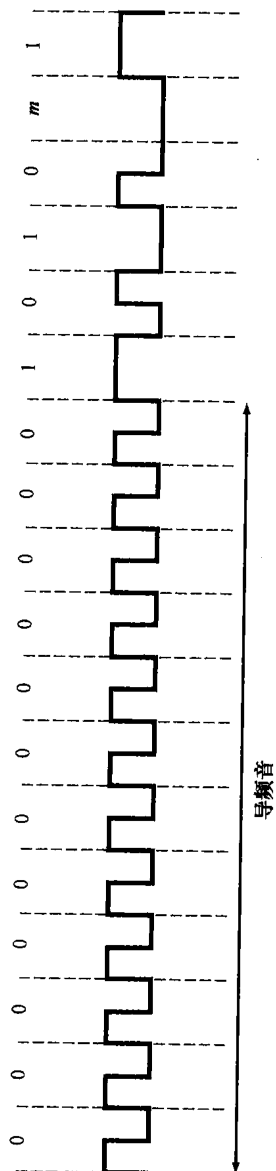


图4-6 有导频音的FM0报头

4.3 标签的状态机

Gen-2 标签的行为可以用一个有限状态机来描述。当需要考虑全部的标志位和计数器时，这个有限状态机需要大量的状态。尽管如此，标签有限状态机可以减少到由七个状态组成的有限状态机。当需要考虑标志位和时隙计数器时，七个状态中的每个状态，又围绕着多个状态。减少后的七个状态的有限状态机需要这些间隔标志位和时隙计数器的相关信息。因此，减少后的七个状态的有限状态机能够用来决定基于一个给定当前设置和当前输入的标签的行为。下节将描述减少后的七个状态的有限状态机中的每个状态。然后，在查询过程期间，将会解释通过标签状态机的移动。该节将得出一个结论，解释在一个访问命令期间，通过标签状态机的移动。

4.3.1 不同标签状态的概述

4.3.1.1 准备状态

准备状态是标签在上电情况下能够进入的两个状态之一。另一个标签上电时可以进入的状态是死亡状态。如果当标签在上电之前是处于死亡状态时，标签将会在上电时进入死亡状态。如果标签先前没有死亡，那它将会在上电后进入准备状态。

处于准备状态的标签，不会参与查询回合。当阅读器发布一个查询命令，将开始查询回合，直到下一个查询命令中止[1]。因此，查询命令指示结束当前的查询和下一个查询的开始。阅读器使用查询回合来获得大量标签的 EPC 号码。

当接收到一个查询命令时，标签将结束准备状态。查询命令包含了标签选择一个随机数载入到时隙计数器的参数。如果随机数是 0，随后标签将会发送它的回复状态。当随机数不为 0 时，标签将会发送仲裁状态。

4.3.1.2 仲裁状态

仲裁状态下的标签将参与当前的查询回合，它们的时隙计数器包含一个非 0 值。在此状态下的标签等待它们的时隙计数器到达 0 为止。当标签的时隙计数器到达 0 时，标签将转变到回复状态。

4.3.1.3 回复状态

回复状态是当标签发送它的 EPC 号码给阅读器时的两个状态之一。回复状态中的标签有一个时隙计数器为 0，并反向散射它们的 RN16（RN16 是一个 16 位的随机数）给阅读器。反向散射 RN16 是标签发送 EPC 号码给阅读器的两个阶段中的第一个阶段。

如果阅读器正确地接收到了标签的 RN16 数，阅读器将会发送一个带有 RN16 的确认（ACK）命令给标签。阅读器接收不到 RN16 的可能原因有：①两个以上的标签在发送 RN16 时发生碰撞；②射频信号干扰了 RN16；③阅读器错过了 RN16 信号。标签随后将会反向散射它的 EPC 号码，还会发送一个 PC（描述了标签的物理

特性）和一个用来进行错误检测的 CRC。接收到带有正确的 RN16 的 ACK 命令使标签从回复状态转到确认状态。

标签只有在接收不到来自阅读器的任何命令时，才会在一个限定的时间内保持在回复状态。在这个时间之后，标签将会自动地回到仲裁状态。

4.3.1.4 确认状态

当标签发送它的 EPC 号给阅读器之后，将会进入确认状态。确认状态是访问命令（读和写命令）的网关状态。但是，标签在确认状态不会死亡。

确认状态与回复状态类似，有一个定时器，可以在每次从阅读器接收到一个命令后复位。如果在特定的时间内，没有从阅读器接收到命令（即定时器超时），标签则会自动地回到仲裁状态。

4.3.1.5 开放状态

开放状态是访问命令专门使用的两个状态之一。在开放状态中没有定时器。因此，如果没有阅读器命令时，只要标签有供电，它将保持开放状态。标签可以从开放状态进入到安全状态。

4.3.1.6 安全状态

标签可以使用一个可选的密码特性，它需要阅读器提供正确的密码，以便于执行任何的访问命令。密码的默认值是 0，这意味着密码保护没有实现，或是没有被激活。设置一个非 0 值的密码，可以使密码具有保护的特性。

如果标签的密码特性被实现和激活，那么所有的访问命令必须从安全状态执行。在标签进入到安全状态之前，必须提供正确的密码。

4.3.1.7 死亡状态

RFID 标签包含识别一个特定物体的信息。ID 号是唯一的，并且与某一特定的物品相关。因此，使用 RFID 时，对于安全和隐私的关注开始上升。这些关注中的一些包含（一些关注）了读 RFID 标签来跟踪人员。小偷读标签来判定哪个顾客携带着贵重的物品。偷听者读药品上的标签来判定某人的疾病状况。

解决这个问题的一个方法是提供关闭标签的能力。这可以参考 Gen-2 中使标签死亡的方法。死亡操作不能被撤销，并且永久性地破坏标签。在死亡后，标签将不会响应任何命令。因此，标签死亡之后，没有任何设备能够读标签的 EPC 号。

在标签接收到正确的死亡密码和死亡命令之后，将会进入到死亡状态。从此开始，标签上电时将会代替准备状态，进入死亡状态。在死亡状态中的标签将不会响应阅读器的任何命令。

4.3.2 查询过程期间通过有限状态机移动的概述

当阅读器想要读在识别范围内标签子集的 EPC 号码时，它将使用上面提及的以选择命令开始的命令。假设全部标签有至少 10s 的能量，而且没有标签会死亡，当上电时，所有的标签将进入准备状态。

首先，选择命令通过更改 SL 和查询标志，来限制阅读器识别范围内的标签到一个特定的状态。标签将会保持在准备状态。

其次，查询命令初始化查询回合，并具体化参与查询回合的标签子集。在响应查询命令时，每个标签将选择一个 Q 位的随机隙号码把值载入到它的时隙计数器内。当标签的时隙计数器到达 0 时，它将会反向散射它的 RN16 给阅读器。选择随机时隙号码为 0 的所有标签将迅速地响应[1]。假设所有的标签与查询匹配，并且标签的时隙计数器不为 0，那么它将进入仲裁状态，不会反向散射它的回复[1]。尽管如此，如果标签的时隙计数器为 0，它将会反向散射它的 RN16，并进入到回复状态[1]。

一个有非 0 时隙计数器的标签，将会保持仲裁状态，直到它的时隙计数器等于 0 为止[1]。在仲裁状态的标签将会响应所有匹配的查询，查询重复和查询调整命令。当时隙计数器到达 0 时，标签将会反向散射它的 RN16，并进入到回复状态。

阅读器可能接收到，也可能接收不到标签反向散射的 RN16。碰撞可能会使 RN16 产生错误，或者阅读器不能够检测到标签的响应。如果阅读器成功地接收到了反向散射的 RN16，它将会使用接收到的 RN16 作为确认命令的 RN 参数，来发送一个确认（ACK）命令。当标签的时隙计数器到达 0 时，标签会储存它反向散射的 RN16，并对它存储的 RN16 与确认命令内的 RN16 进行比较。如果它们匹配，那么标签将会发送它的 EPC 号码给阅读器；否则标签则会忽略这个命令[1]。当标签发送它的 EPC 号码给阅读器时，它将会进入到确认状态。确认状态是网关状态，是标签状态机覆盖访问（读、写、安全）命令的一部分。

4.3.3　在一个访问命令期间，通过标签状态机移动的概述

访问命令提供阅读器运行高级函数的能力，这些函数包括读写数据，改变密码，锁定标签上的存储器，禁用标签或是使标签死亡等。访问命令可以用来读写标签附着物品的相关信息，这个标签允许信息跟随物品移动或者驻扎。使用这个特性的其中一个例子是需要写入温度的易腐蚀货物区域，或是其他环境状态中允许接收机来决定在存储和运输期间，物品是否保持在合适的环境状况下。这有助于阻止坏的货物流入市场，对健康造成损害。另外，标签的 EPC 号码能够存储在中心数据库，相关物品的信息则能够存储在数据库内。

在标签响应任何访问命令之前，标签必须在确认的状态下。这意味着在标签的 EPC 号码被识读之后，标签能够被迅速地访问。因此，在发布任何访问命令给标签的第一步是读标签的 EPC 号码。（把标签置于确认状态）4.3.2 节描述了这个过程。

在读了 EPC 号码之后，阅读器必须发送带有 RN16（在确认命令来读 EPC 号码时使用的 16 位随机数）的 Req_RN 命令，以便于标签进入到开放状态，或是安全状态。从确认状态后进入哪个状态，取决于访问密码的值。一个全 0 的访问密码意

味着访问密码特性没有实现或是没有激活[1]。如果访问密码没有实现或是没有激活（访问密码为0），标签会进入安全状态来响应 Req_RN 命令。如果访问密码非0，那么访问密码特性已经激活，标签将进入到开放状态[1]。在响应有效的 Req_RN 命令（带有正确访问密码或是访问密码为0的正确的 RN16）时，标签将会反向散射回一个在其他访问命令中使用的句柄[1]。当标签在开放状态时，并且有一个非0的访问密码，访问密码必须与正确的句柄和标签转换到安全状态的访问密码一起使用。

锁定命令允许锁定存储位置，或者使存储器只读或只写，或者是同时可以读写操作[1]。如果存储器位置先前被锁定，锁定命令可以解锁存储器的位置，使它们可读、可写或者是可读写[1]。只有在标签处于安全状态时，锁定操作才有效[1]。其他访问命令，包括死亡命令，能够在标签处于开放状态时执行。

总之，在执行一个访问命令时，有三个步骤。首先，如4.3.2节所述的，必须先读取标签的 EPC 号码；第二，Req_RN 命令必须与正确的 RN16 一起发送给标签；第三，如果访问密码为非0时，标签必须使用访问命令来进入到安全状态。此时，包括锁定命令在内的任何访问命令都能够被执行。

4.4　标签查询特性

Gen-2 协议定义了一个命令集合。命令的长度是可变的。因此，Gen-2 协议试图最小化这些最经常使用的命令的长度。

Gen-2 命令最常用来在给定的时间内，对阅读器范围内的标签的信息进行采集。这个过程被称为清查标签，并且这些用来清查标签命令的长度是最小化的。Gen-2 使用一个时隙 Aloha 碰撞避免算法来减少一个碰撞产生后数据错误的数量。在时隙 Aloha 通信方案中，时间被分为一定数量离散的时间片，这称为时隙。想要发送数据的设备，会在一个时隙的开始处开始发送数据。此外，传输不能在相邻的时隙间重叠。因此，当有碰撞时，只有一个时隙的数据是损坏的。

没有了时隙的限制，在设备 B 传送数据期间，设备 A 开始传送数据，并且直到设备 C 传送数据时，设备 A 也仍然在传送数据，这都是有可能的，而在这个过程中，设备 B 和设备 C 传送数据的过程是不重叠的。在这种情况下，三个消息将会丢失，因为在时隙 Aloha 环境中，只有一个消息可以通过。

4.4.1　查询命令概述

在 Gen-2 协议中，选择标签唯一的 EPC 号码是重要的操作之一。有四个命令用来查询标签：①选择命令、②查询命令、③查询重复命令、④查询调节命令。每个命令在查询期间，都执行不同的操作。

4.4.1.1　查询

查询命令指示 Gen-2 查询回合的开始[1]。查询命令与阅读器向标签报头一起，定义全部的标签向阅读器链路的特性。相邻标签使用查询命令中的三个区域连同阅读器向标签的报头，来决定标签向阅读器通信链路的参数。这三个区域称为 DR、M 和 TRext 区。如 4.2.4 节和式（4-4）描述的，标签使用 DR 区来计算标签向阅读器链路的数据速率。M 区用来选择 FM0 或者三个密勒编码方案之一，来编码标签反向散射的数据。查询命令的 TRext 区决定在响应时标签是否包含标签报头。TRext 区是一位的，当 TRext 为 1 时，将包括标签报头；当 TRext 为 0 时，将不包括标签报头[1]。

当标签接收到一个查询命令时，标签首先检查查询命令提出的标志值是否与标签存储的标志值匹配。只有标签存储的全部标志值与查询命令的标志值匹配时，标签才会参与查询回合，否则，标签将会忽略查询命令。当标签参与查询回合时，标签将会挑选一个长度由查询命令的 Q 区定义的随机数。Q 区的值在 0 ~ 15 的范围内，并且可以决定所有标签可以使用的时隙数量。时隙的数量 N_{slots}，在给定 Q 值情况下，由式（4-5）定义：

$$N_{slots}=2^{Q}-1 \tag{4-5}$$

标签随后会等待那个时隙，并通过发送 RN16 给阅读器来响应。RN16 是一个 16 位的随机数，用来访问一个唯一的标签，因为两个标签同时选择一个相同 RN16 的可能性是极小的。参数 Q 说明它的 16 位随机数产生器使用了多少位，供它的时隙来表示时隙计数器[1]。阅读器使用 RN16 值来读标签唯一的 EPC 号码，并作为读、写和安全这样高级别功能的一个句柄。查询命令的长度为 22 位。

4.4.1.2　查询重复

查询重复命令通过使所有参与的不能读标签的时隙计数器减 1 来提前下一个时隙。选择该时隙来回复的任何标签将发送它们的 RN16 值给阅读器。查询重复命令可能是 Gen-2 命令中使用最频繁的命令，因此也是 Gen-2 命令中最短的命令，它只有四位的长度。

4.4.1.3　查询调节命令

查询命令总是开始一个新的查询回合，但是有时候，阅读器只需要稍微调节一下作为时隙数量的 Q 值，或者是要求所有非查询标签来选择一个响应的新的时隙数量。查询调节命令提供了一个机制，使 Q 值加 1、减 1，或者保持 Q 值一定。在上述三种情况中，没有查询到达的标签，即标签没有反向散射回它的 EPC 序号，将会选择新的时隙来进行响应。查询调节命令的长度比查询命令短，只有九位，并且不会开始一个新的查询回合。开始一个新的查询回合对被查询标签的当前状态会有额外的副作用，这对完成查询过程所需的时间产生不利的影响。

4.4.1.4　选择

选择命令定义了参与即将进行的查询回合标签的数量。虽然阅读器范围内的所

有标签都可以参加当前的查询回合，但这是没有必要的。因此，只有在阅读器范围内的一个标签子集会参加当前的查询。选择命令提供了一系列的特性来通知标签它们是否需要参加当前的查询回合。描述用选择命令来定义参与当前查询回合标签的子集的例子，将在本章随后的部分提供。

4.4.2 会话的使用

当两个阅读器在同一个时间帧试图与一个标签通信时，将会产生一系列的问题。在给标签写数据和从标签读数据或是EPC序号时，上述问题是真实存在的。RFID系统必须允许多个阅读器与单个标签进行通信。Gen-2标签标准提供了一些有限的支持，它允许在同一时间帧内，四个阅读器与一个标签进行通信。这个机制在Gen-2内被称为会话。

在Gen-2中定义了四个会话，每个会话都有它自己的查询标志[1]。剩下的标签状态变量将共享全部四个会话[1]。会话被标记为S0、S1、S2和S3[1]。这四个会话能够使四个阅读器独立地与一个标签通信，来读取此标签的EPC号。查询标志的初始值取决于阅读器使用的会话和查询标志的先前值。如果在持续的时间过期之前，阅读器通过电磁波辐射供给标签能量，使标签上电后，会话S0的查询标志总是设置为A，而会话S1、S2和S3的查询标志设置为A或B的先前值，否则查询标志会将默认为A[1]。

由于标签没有板上电源供应，每个会话有一个相关的持续时间，这段时间是每个会话查询标志在阅读器停止传输能量给标签后保持的时间长度。标签能够使用一个电容来存储少量的能量，存储在电容上能量的多少，决定了持续时间。因此，每个会话有它自己的电容，并且每个电容的值不相同，这使得电容存储能量的大小也不相同，从而决定了查询标志持续的时间。当标签在会话S2和S3时失去了能量，查询标志的持续时间最短为2s，而会话S1的持续时间在500ms~5s之间，会话S0没有持续时间，即持续时间为0[1]。在持续时间过期之后，即便标签仍有电源供应，S1的查询标志也将会还原成A[1]。

会话是很有用的，因为它们能够用来把所有标签的查询标志设置成一个已知的值。然后，当阅读器读每个标签的EPC序号时，标签的查询标志会转换，这个标签将不再参与查询回合。因此，当不再有标签存在给定的查询标志时，阅读器可以确认在给定的范围内它已经识读了所有的标签。这需要阅读器选择适当的会话来使用。当S2和S3都有一个持续的时间，并且只要在持续时间过期之前阅读器能够给标签提供能量，那么上述过程将会工作。会话S0没有持续时间，因此在这种情况下，阅读器必须持续的给标签提供能量。这是通过阅读器在发送命令之间发送载波（CW）来实现的。使用会话S0时，在这个过程期间任何失去能量的标签的查询标志不管它的先前值多少，都将会还原成A。假设阅读器正在查找查询标志为A的标签，由于这个问题，阅读器可能会多次读同一个标签的EPC号。当使用S1时，必

须保证整个查询回合的时间比持续时间短。如果查询回合时间比持续时间长，那么不管先前的值是多少，查询标志都将会还原成 A。当使用会话 S1 时，如果持续时间过期，标签连续进入这个过程是有可能的。因此，会话 S1 只能在短的查询回合中使用。尽管如此，在命令之间保持载波（CW）可以阻止由于持续时间到期引起的查询标志的复位。

会话提供了一些有限的支持，可以使四个不同的阅读器来访问同一个标签，并在相同的期间内读取它的 EPC 序号。为了实现上述过程，每个阅读器必须选择不同的会话来避免每次其中一个阅读器读 EPC 序号时，连续地在查询标志 A 和 B 之间转换。由于更高级别的读、写、访问功能不仅仅只需要查询标志，会话的使用不能提供多个阅读器执行高级别命令的能力。但是，这是有利的，因为当多个阅读器对同一个标签执行高级别的操作时，例如数据一致性和状态一致性等问题将会产生。

4.4.3 选择命令的特性

选择命令用来选择阅读器范围内标签的子集，这个子集将会参与即将进行的查询回合。这个子集可能包括阅读器范围内全部的标签，可能是一部分标签，也可能不包括任何标签。开始一个查询回合的查询命令使用选择标志（SL 标志）和四个查询标志中的一个来选择参与查询回合理想的标签子集。选择命令允许阅读器改变已经选择的标志（SL 标志）或者是存储在标签内的四个查询标志中的一个。

选择命令包含四个强制区域和一个可以识别参与即将进行查询回合的标签子集的可选区域。第一个强制区域是目标区域，它允许阅读器改变 SL 标志，或者是四个会话（S0、S1、S2 和 S3）中的一个查询标志，但是不改变 SL 或者查询标志[1]。第二个强制区域是动作区域，它可以具体化所有匹配选择命令参数的标签将参与的动作。可用的动作可以设置 SL 标志的值，或者把查询标志的值设置为 A 或是 B[1]。只有在标签的参数与选择命令在 MemBank、指针、长度和 Mask 区域的一系列参数匹配时，这个动作才会进行[1]。MemBank、指针、长度和 Mask 区域用来进一步指定标签超过 SL 或者是查询标志时的匹配标准。Mask 区域是可选的，并且也有一个在 0 ~ 255 字节间的可变长度[1]。当使用 Mask 区域时，它将包含一个 N 字节的字符串，字符串必须与参加下一个查询回合的标签存储器内具体位置内的内容（比如 EPC 序号）匹配。长度区域定义为 Mask 区域的字节长度，用字节表示[1]。Mask 能与标签存储器四个类型中的一个进行比较。Gen-2 把标签存储器空间分成 4 个部分：①第一部分是保留存储器 bank，包含了标签相关的密码；②第二部分是包含 EPC 号的 EPC 存储器 bank；③第三部分是 TID 存储器 bank，包含了标签和供应商的具体信息；④第四部分是包含了用户定义数据的用户存储器 bank[1]。MemBank 区域指定了 Mask 将会与这四个存储器 bank 中的哪个进行比较[1]。指针区域定义了 Mask 进行比较的存储器开始地址[1]。指针提供了一个存储器地址来存储由 Mem-

Bank 区域指定的相关存储器 bank[1]。复杂的和多个子集的标签通过发布多个选择命令来进一步定义和提炼标签子集来得到便利。

一个截短区域的第七区域，是选择命令的一部分。截短区域与 Mask 区域结合使用，来减少标签响应 ACK 命令后反向散射回阅读器的 ECP 位的数量[1]。只有当 Membank 指向 EPC 存储器或是 Mask 被指定，或是那些目标制定的 SL 标志时，截短命令才会使用，这个特殊的选择命令是查询回合开始之前的最后一个选择命令[1]。当截短命令被激活时，标签在紧跟 Mask 之后，只发送紧跟 Mask 之后的 EPC 序号的一部分[1]。因此，截短命令能够用来缩短标签回复 ACK 命令时的长度。截短区域的使用在 1.6 节部分也进行过探讨。

4.4.4 查询命令的特性

查询命令开始一个新的查询回合，识别标签中哪个子集将会参与新的查询回合，并选择标签到阅读器的数据编码和数据速率[1]。选择命令用来更改 SL 标志和标签的四个查询标志。查询命令指定 SL 标志的一个值、一个特殊会话和这个特殊会话查询过程的一个值[1]。特殊会话的 SL 标志和查询标志与查询命令中的请求值相匹配的标签，将会参与新的查询回合。

查询命令包含七个区域和一个进行错误检测的 CRC，这七个区域中的三个，DR、M 和 TRext 区域分别用来定义标签到阅读器的数据速率和数据编码，以及导频音是否包含在标签到阅读器的报头[1]。在 4.4.1 节讨论了 DR 和 M 区域。TRext 区域说明了标签响应阅读器时，在反向散射标签到阅读器报头之前，是否反向散射导频音[1]。

其余的三个区域，即选择、会话和目标区域定义了标签决定是否参与查询回合时需要的 SL 标志和查询标志参数[1]。选择区域指定 SL 标志的值，即这个标签必须参与查询回合。选择区域可以指定 SL 标签声明或者是不声明[1]。会话区域定义了 S0、S1、S2 和 S3 这 4 个会话中的一个[1]。查询目标区域把查询标志的值定义为 A 或者是 B，这个会话由会话区域指定[1]。使用这 3 个值，每个标签可以决定它是否应当参加查询回合。所有的会话区域指定的会话 SL 标志和查询标志值与查询命令指定的值相匹配的标签，将参与查询回合，那些具有不同标志值的标签将不参与查询回合。

4.4.5 查询重复命令的特性

查询重复命令是 Gen-2 协议中最短的命令。查询重复命令指导全部的标签推进到下一个时隙。在响应查询重复命令后，每个标签减少它的时隙计数器值。当标签计数器值到达 0（或者选择 0 作为随机 Q 位数）时，它将会响应，并反向散射回它的 RN16 给阅读器。

查询重复命令有一个参数，会话区域，它可以用来预示标签的哪个会话需要减

少它的时隙计数器值[1]。如果一个标签在给定的会话 S0 中参与查询回合，并且接收到另一个会话的查询（S1、S2、S3）命令，它将忽略查询重复命令[1]。

4.4.6 查询调节命令的特性

查询调节命令用来调节 Q 值，然后选择一个新的时隙计数器，或者在没有改变 Q 的情况下，指导标签来选择一个新的时隙计数器。查询调节命令有两个参数：会话参数和 UpDn 参数[1]。会话参数指示查询命令与哪个会话相关[1]。和查询重复命令一样，在一个会话中，参与一个查询回合的标签将会忽视具有不同会话参数的全部查询调节命令[1]。UpDn 参数定义了标签响应查询调节命令的动作。

查询调节命令能够使标签采取三个动作。第一个动作是保持先前的 Q 值，然后选择一个新的 Q 位随机时隙号码[1]。第二个动作用来对先前的 Q 值增加 1，然后选择一个 Q 位随机时隙号码[1]。第三个动作用来对先前的 Q 值减少 1，然后选择一个 Q 位随机时隙号码[1]。

4.5 标签单一化

标签单一化是一个过程，它可以为读 EPC 号码，或者是在大量标签中的单个标签的高级命令（读、写或者安全命令）建立通信。单一化一个标签取决于当前的操作环境和标签数量的组成。虽然大多数方法可以在全部的情况下工作，但是选择一个需要最少时间量的方法往往是有利的。这有助于在更大系统中避免 RFID 成为瓶颈或者至少减缓这种瓶颈所带来的影响。

4.5.1 EPC Gen-2 标签数据编码分类

EPCglobal（全球电子产品码）组织开发并维护 Gen-2 标准，而且定义了一个 EPC 号编码集。这些 EPC 编码类型是由 EPCglobal 维护的标签数据标准描述的[2]。当被询问包含定义了标签数据编码和 Gen-2 字符单元中标签 EPC 序号长度的信息的 EPC 序号时，标签将反向散射协议控制（PC）位。这个标准定义了 11 个不同的数据编码 EPC 标签[2]。在一个 Gen-2 标签中 EPC 存储器 Bank 的最大长度是 496 位（31 个 Gen-2 字）[1]。EPC 序号的理论最小长度为一个 Gen-2 字，但是标签数据标准定义的最短 EPC 号码是六个 Gen-2 字或者 96 位[2]。一个 Gen-2 字等于 16 位[1]。更短的编码（64 位）保留了头部，但是只包含了与旧标签的向后兼容。最终这些头部将会被新的编码回收利用。

11 个不同标签数据格式都包含了不同格式的不同信息或相同信息。编码格式的其中一个是通用识别器（GID-96），它是一个基本的 96 位 EPC 号[2]。GID-96 包含了四个区域：①头部识别，即这个 EPC 号码是根据 GID-96 的说明进行编码的；②通用管理者号码识别资产的拥有者；③对象类识别资产的类型；④序列号识别一

个资产的一个独一无二的示例[2]。对象类和序列号是被拥有者指派的，并且在拥有者的号码集当中是独一无二的[2]。

GID-96数据编码标准与IP地址分配非常相似，大的组织将被给予包含独立IP地址号码的B类地址空间，并且把这些IP地址分配给它们的内部组织。对象分类可以用来对组织进行分裂，可以分配一个C类地址空间来使用这个特定的分裂。序列号可以被认为是在一个办公室内被分配给一个特定设备的单个IP地址。

4.5.2　选择单个标签

选择单个标签需要阅读器区域内出现的标签集的大部分信息。可以使用标签EPC序号作为选择命令的Mask区域来选择单个标签。在选择命令之后，此时被选择标签将声明具有与发送的选择命令的mask区域相同的EPC序号。选择单个标签是执行标签上更高级的读和写命令的捷径。

查询命令的发布用来查询那些选择标志已经声明，并且Q值可以设置为0的标签，从而使单个选择标签响应。一旦标签响应，并得到确认，更高级别的命令将能够执行。

4.5.3　选择一组标签

标签组可以基于它们SL标志的值、特定会话的查询标志或者基于它们存储器的内容（即EPC序号）被选择。SL和查询标志不提供选择标签正确组的保证。不能保证子集中想要的标签都会出现在子集中，也有可能不想要的标签也会包括在子集中。

尽管如此，基于存储器内容的选择命令有把标签放入子集的能力，能够用来创建子集。通过使用选择命令的掩码特性，这是有可能的，即需要标签的指定存储器数据与掩码区域匹配，从而通过选择命令采取相应的动作[1]。EPC序号有一个结构，可以用来标准化这个过程。因此，子集中想要的标签能够被放入到子集中，不参与查询回合的标签也会进入子集。

在4.5.1节描述的GID-96编码结构将在本章内使用，EPC序号的编码用来描述本章出现的概念。提出的方法也可以在其他编码中使用。GID-96包含了四个区域，其中头部区域不会因GID-96类型编码而改变。第二个和第三个区域，即通过管理号和序列号区域分别用于指明资产的拥有者和资产的类型。第四个区域是序列号区域，在一个给定的一般管理者代码和目标类结合中是独一无二的。因此，一般管理者代码能够用来选择基于它们拥有者或者制造商的一组标签。这个子集能够通过使用目标类在给定一般管理者代码中选择一个特定的资产类型来进一步提炼。因为多个选择的命令能够连在一起，使得在一组不同的拥有者中选择一组标签类型成为可能。

4.5.4 选择全部的标签

有时，全部标签参与查询回合也是有利的。这可以通过发布一个选择命令来得到保证。为了实现上述目的，需要使用会话 S0。选择命令将更改会话 S0 的查询标志，使用“000”作为目标区域，而且不使用掩码或者截断功能[1]。这会使匹配第一个选择命令的全部标签设置它们的会话 S0 查询标志为 A，并且其他不匹配的标签设置 S0 查询标志为 A[1]。在这点，全部的标签将设置它们的会话 S0 查询标志为 A，因为在选择命令中没有指定的 Mask（掩码长度为 0），这意味着所有的标签匹配[1]。通过发起一个查询命令，它的 Sel 区域全部设置成“00”或“01”值，查询区域设置为会话 S0，目标区域设置为 A，查询回合就会开始[1]。现在所有的标签将参与这个查询回合。

4.6 权衡

零售物品是 Gen-2 协议 RFID 标签使用的首要目标。包含一定数量产品的货盘会发送给顾客。Gen-2 标签会被频繁地安装到货盘的物品上。虽然条形码需要视线距离来读取信息，Gen-2 协议的 RFID 标签可以通过其他物体来读取。因此，当一个货盘到达时，它必须得打开。但是，如果物品被安装上 RFID 标签时，可以通过 RFID 在不打开货盘的情况下得到货盘内信息，从而节省了大量的时间。

在 Gen-2 协议中有一定数量的权衡。第一个权衡是参与给定查询回合的标签子集的建立。这个子集的创建是选择和查询命令的结果。这个部分将描述这个权衡的更多细节，并给出了不同标签子集有用的例子。选择正确数值优化的集合，将减少 RFID 系统读货盘上所有 RFID 标签的时间，并增加了吞吐量。

另外一个权衡是在阅读器到标签、标签到阅读器、阅读器数据编码和标签数据编码的选择中。这些值通常取决于系统运行的射频环境的特性。例如，在一个非常低的射频噪声的环境中，快数据速率和快（简单）数据编码将会更适合产生更快的阅读器标签交互。相反地，在具有显著噪声、较低数据速率和更高健壮性数据编码的环境中，例如密勒 4 或者密勒 8，将有利于减少位错误（增加信息正确解码的可能性），但是这些数据编码导致了一个较低的数据速率。

在本部分给出的例子中，假设使用通用的 96 位 EPC 序号（GID-96）格式。此外，假设已经持续至少 10s，因为货盘上的任何标签已经上电。

4.6.1 查询货盘上包含一种类型产品的标签

最简单的情况是，货盘上只包含了一个生产商的一种类型的物品。需要记住的是，选择命令包含了一个能够用来只把那些存储器与掩码匹配的标签放入参与当前查询回合的 Mask 区域。也需要记住的是，截短区域能够使标签发送回一个缩短的

EPC 序号。在这种情况下，头部、一般管理者代码和目标类值对于所有的标签将会相同，只有序列号将会不同。

在这种情况下，全部标签必须包含在参与查询回合的标签子集中。这可以通过发布两个选择命令来保证。第一个选择命令将定位所有 SL 标志声明的标签，并使用动作码“110”，这不会造成匹配标签的改变，使其他不匹配的标签声明它们的 SL 标志。第二个选择命令将用它们声明的 SL 标志，使用不改变的“110”动作来定位全部（这是所有标签的数量）。这个 Mask 将用来设置头部、一般管理者代码和目标类。利用声明的截短标志，每个标签将只返回它的序列号。这个序列号长度为 36 位，而整个 EPC 序号的长度为 96 位[2]。这节省了 60 个字节或者在读每个标签时反向散射的 62.5% 的减少量。从这点看出，阅读器简单地用查询命令开始一个查询回合，并用查询、查询回复和查询调节命令连续地读标签的 EPC 序号，直到所有的标签被识读为止。

当头部，一般管理者代码和目标类的值在先前的时间没有知道时，它们可以通过从单个标签内读一个 EPC 序号得到这些值。一旦从单个标签 EPC 序号获得头部、一般管理者代码和目标类的值之后，上述描述的过程可以用来读剩下的标签。

4.6.2 访问货盘上包含一种类型产品的标签

如 4.3.3 节描述的，当标签的 EPC 序号被读取时，或者全部标签被读后，访问一个标签能够进行。为了在全部标签被读取后访问一个标签，需要使用 Mask 设置为那个标签整个 EPC 序号的选择命令。然后，发布一个 Q 设置为 0 的查询（一个单一的时隙）。标签将会在单一的时隙内响应，阅读器能够确认标签，并访问标签。如果标签在被读之后没有立刻被访问，等待直到全部被读的标签使用上述方法访问了那个标签是很重要的。这是由于上述方法可能改变标签的 SL 标志，会造成在当前查询回合期间内确保全部剩下的标签已被识读时出现问题。

4.6.3 查询货盘上包含一个单一生产商多种类型的产品的标签

在这种情况下，货盘包含了一定数量同一生产商的不同类型的物品。因此，所有标签的头部、一般管理者代码将会相同，但是目标类和序列号是不同的。4.6.1 节描述了与这一个情况相似的过程，只是在第二个选择命令的 Mask 域中仅仅指定头部和一般管理者代码。如果在货盘上只有少量不同类型的产品或者托盘上很大百分比是单一的一种产品，那么这可能对在 Mask 中包含目标类以减少标签反向散射给阅读器的 EPC 的量是有利的。尽管如此，必须仔细平衡这个权衡，因为必须为每种类型产品发起两个选择命令和多个查询（查询、查询重复和查询调节）命令。随着每个类型产品的物品数量的减少，由于在查询回合中增加了设置和运行的消耗，标签以一个较短的 EPC 序号响应阅读器的优点将不复存在。

4.6.4 访问货盘上包含单一生产商多个产品类型的标签

在这种情况中，标签能够根据使用目标类的产品类型分组成更小的子群。因此，为了找到想要的标签，只需要读取一个特定的标签子集。因为子群包含了比查询整个货盘更少的标签，这可以减少访问特定标签所需要的时间。在正被读取或者全部子群中的标签（具有特定目标类区域的标签）被读取之后，标签可以迅速地被访问。在 4.6.3 节提到的发送一连串选择命令来建立子群开销的权衡必须与读货盘上全部标签需要的时间进行比较。

4.6.5 查询货盘上包含多个生产商的多个类型产品的标签

在这种情况下，托盘包含了不同生产商的多种类型的产品。因此，全部标签只有头部部分 EPC 号是相同的。头部区域的长度是 8 位，它指定了标签上数据编码的类型[2]。选择命令的掩码和截短特性能够用来减少发送回的 EPC 号的数量，但是使用标签散射回的 GID-96 编码只保留 8 位或者是 EPC 信息的 8.33%。一般管理者代码、目标类和序列号的值在这个货盘类型中将会改变。

因此，使用选择命令的掩码和截短特性必须与发布的额外选择命令的额外支出相平衡。如果相同物品的数量比较少或者货盘内包含大量不同生产商的物品，那么发布额外选择命令的消耗可能会增加查询操作所需要的时间。在这种情况下，仅使用头部来使用选择命令的掩码和截短特性是非常有利的，并同时从全部标签读取 EPC 序号。

4.6.6 访问一个货盘包含的多个生产商的多个类型产品的标签

在这种情况下，可以基于生产商和产品类型，分别使用一般管理者代码和目标类区域分离标签。如果这些值是已知的，那么子群将容易形成，并排除托盘上很大一部分的标签。在标签被读或者通过管理号和目标类区域匹配的全部标签被读取之后，想要的标签可能迅速地得到访问。建立子集所需的选择命令序列增加的开销和读托盘上所有标签的时间的权衡，必须要评估。因为如果货盘上只有少量同一类型和同一生产商的物品，可以较快速地形成子群。

4.7 开放问题

Gen-2 协议效率的开放问题包括决定最好的数据编码和数据速率的物理层设置，来减少位错误以及在优化查询和访问过程最小化标签上执行操作需要的时间。

最好的物理层特性的确定包括权衡更加健壮的通信链路来较少位错误，以及由此造成的数据速率的降低。预测将来射频环境状况的易处理的和精确的模型是非常重要的，但是很难规划。对于广泛的商业部署，这些模型必须适应一个宽范围的运

行状况，例如工厂，仓库或者是移动的卡车。而且，射频信号通过物品的能力也是重要的，因为当信号穿过物品时，如果信号衰减地过快，阅读器或者标签将不能正常地接收到信息。

最小化读全部标签 EPC 序号的时间需要基于查询标签集组成和每个标签状况的预测。知道一个被查询货盘的类型，就能够使阅读器选择使用哪个 Mask 集来减少每个标签反向散射的 EPC 序号的大小。尽管如此，这些信息可能并不有效，并且预测或是游戏理论类型模型必须得到发展，以便于快速地决定货盘的类型，或者这些信息必须由货盘来提供给阅读器。

4.8 结论和未来研究方向

Gen-2 协议的特性使用户能够对一个群组内的标签交互时间进行优化。在零售和分销领域减少一个货盘物品的查询时间从而增加吞吐量，这是非常有用的。但是，仍然有大量的开放性问题，如能够更加抵抗出错的物理层和用于决定如何与标签自身通信的模型和方法学。

减少射频环境的干扰或者环境噪声造成的位错误，可以减少货盘上精确查询所需要的时间。信号处理技术可以为破译由于噪声造成的信息乱码的正确数据提供帮助。而且，有趣的是在碰撞期间可以识别至少一个从标签反向散射回的 RN16。这将极大地改善在给定期间内能够被识别的标签数量。

由于其他物品造成的射频信号衰减，使阅读器与标签通信的范围减小。这限制了货盘的大小，因为货盘需要保证所有的标签都能够被精确地识读。改变协议的物理层的特性是不明智的。因为这将会使基于 Gen-2 协议存在的产品过时。

使用正确的混合命令来查询一组标签是非常重要的。虽然简单地尝试在相同时间内，在一个范围内查询全部标签也可以工作，但是选择命令的截短和掩码特性的益处将使用不到。当附加时间来定义将参与查询回合的标签子集时，适当地使用选择标签是非常重要的。准确地预测被查询货盘的类型可以最小化查询的时间。而且，选择合适的查询命令的混合可以最小化读全部 EPC 号的时间。研究这两个领域的模型或者政策是必要的。

参 考 文 献

1. EPCTM radio-frequency identify protocols class-1 generation-2 UHF RFID protocol for communications at 860 MHz–960 MHz version 1.1.0, EPCglobal Inc., 2005.
2. EPCglobal tag data standards version 1.3.1, EPCglobal Inc., 2007.

第5章 RFID的认证和隐私

RFID标签是便宜、简单的设备，它能够存储唯一的识别信息，并且能运行简单的计算来保持对数据包较好的查询。这个特性与条形码相比，有一个显著的优点，它允许标签在例如库存跟踪、供应链管理、防盗等不同的领域使用。尽管如此，不像条形码，标签具有更长的扫描范围，可能会被没有认证的恶意阅读器扫描，和遭到包括复制在内的各种攻击。因此，需要一种RFID的安全协议来确保在每个标签和阅读器之间的隐私和认证。本章将概述近年提出的各种安全方法。这些方法包括用来保护RFID标签和标签自身的低计算算法协议的分离设备，其中的两个方法是由本章的作者提出的。本章最后讨论了RFID安全未来的发展方向和关于这个领域的一些开放性的研究问题。

5.1 概述

RFID标签是一个较小的电子器件，被用来识别和跟踪物品。RFID标签可以应用到库存跟踪、供应链管理、防盗等不同的领域。一个RFID系统由RFID标签、RFID阅读器和后台数据库组成。一个RFID阅读器包含了一个RFID发射器和一个RFID接收器，一个控制单元和一个存储单元。这些设备一起工作，来转发和接收在无线电波内存储的信息，这个无线电波是阅读器和RFID标签之间的无线电波。这些信息与后台数据库存储的条目进行交互，而一些阅读器能够访问这个后台数据库。根据标签的类型，它们有能力执行不同的功能，来处理从阅读器传输的信息。

RFID标签有3个分类：被动、半被动和主动。被动标签，即无源标签，是由阅读器的信号来提供能量的，并且只能在数米的小范围内工作。主动标签，即有源标签，使用电池来保持它们的内部状态和能量传输。半被动标签是电池辅助的标签，它使用电池能量来保持内部的非易失存储器，但是仍旧依赖阅读器的信号提供的能量进行数据的收发。它们能够初始化通信，并能在更长的范围内工作，但是相对无源标签，它们的成本也会更高，体积会更笨重。尽管如此，被动标签的使用更加广泛，也更加廉价，如前所述，这使得被动标签能够在更广的范围内得到应用。因此，本章只集中介绍关于被动标签设计相关的设备和协议。

RFID标签能够唯一地识别一种产品类型的物品，而条形码只能识别每个产品的类型。当每个物品的交易历史需要被保持或当一个物品需要被跟踪时，RFID标签显得非常有用。此外，RFID标签不需要视距通信，这可以显著地提高标签的扫描过程。由于众多的优点，RFID标签变得越来越受欢迎，并且在不久的将来，有

可能替代现在的条形码技术。但是，依然存在着消费者隐私保护和其他安全漏洞等问题，这些问题使得RFID标签很容易成为恶意攻击的目标。被动的RFID标签目前的形式很容易受各种类型的攻击，因此，在大范围部署之前，需要使这项技术能够更加安全。所以，隐私和认证是RFID技术需要关注的两个主要的安全问题。

对于RFID标签隐私的两个主要关注点是隐蔽跟踪和查询[1]。隐蔽跟踪处理一个附近的RFID阅读器能够扫描任何标签的问题，因为这些标签会不加慎重考虑地响应阅读器。另外，隐蔽查询是一种从标签采集敏感信息的方法，可以得到一个组织查询的相关信息。一个称为EPCGlobal的组织[2]管理着电子标签码（EPC），这个RFID标签内的码相当于用来存储信息的条形码。兼容EPC的RFID标签有存储生产商和产品码的区域，这使得标签很容易就可以遵循如商店[1]或者企业员工ID号的分配的查询模式。

RFID隐私是生活中众多领域关心的一个问题。下面有几个例子。自动支付收发器，放置在风窗玻璃角落的饰板上，是全世界普遍使用的。一些图书馆实现了RFID系统，使得图书的检出和查询控制更加便利，减少了图书馆员工重复性的劳动压力。部分程度上由美国爱国者法刺激产生的选书的监视问题，激起了关于RFID的隐私问题[4]。最后，国际民航组织（ICAO）颁布了使用RFID的护照和其他旅行文档的准则[5,6]。美国已经授权采用这些标准，作为27个免签证国际公民入境的条件。由于技术挑战和技术参数的改变，这个授权已经延期，有部分原因是因为隐私提倡者的游说。有人可能会认为，信息的验证如果存储在护照上，也同样会成为问题。这带来了RFID如认证方面的安全威胁。

认证是RFID标签另一个主要的安全问题。隐私处理被攻击阅读器所篡改的认证标签，而认证用来处理被虚假标签所迷惑的有效的阅读器。例如，当扫描仿制标签时，认证就发挥了作用。已经证明的是，是可以重写一个标签发给另一个标签的信息，并有效地进行复制[1]。因此，认证和隐私是一样重要的。

给被动RFID标签提供安全机制的重要挑战是被动标签被设计成为低成本的设备，它只有极弱的计算能力[7]。已经提出了许多方案来解决RFID标签的这两个安全威胁。这些方案包括分离用来保护RFID标签和发展标签自身低计算算法协议的设备。本章将讨论其中的一些协议的优缺点，并探讨RFID安全现今的状态和未来的发展方向。

5.2 重要的RFID认证和隐私协议

为了实现向客户的RFID标签提供安全的需求，研究者首先在扫描时禁用标签。换句话说，一旦一件产品通过了检查过程，一个像磁铁的设备将会打乱标签内写的数据，并禁止它再次使用。这个过程现在已经在许多书店和其他零售环境下使用，来防止小偷。但是，这个方法并不是对所有的物品都适合。因此，研究者开始

借用知名的加密方法，不仅发布隐私，也提供包含例如识别证件和电子护照的更高安全风险的标签的彼此认证。这部分描述了禁用标签的缺点，并深入研究了加密方法将如何与 RFID 标签进行比较。

5.2.1 标签死亡协议

第一个用来处理消费者隐私的方法是由 EPCGlobal 公司提出的，这个公司将会监督条形码到 RFID 的转换。它们的方法是“杀死”标签[2]。换句话说，将使标签无法继续工作，使得标签不被恶意的阅读器扫描。这个过程通过阅读器发送一个特殊的“杀死”命令给标签（包含一个短 8 位的密码）来完成。例如，当你推着超市购物车通过一个自动检查器并付完款之后，购物车内所有相关的标签都将被“杀死”。

虽然杀死标签可以处理消费者的隐私，但这也取消了消费者所有售后的好处。这些售后好处类型的其中一个例子是能与称为智能机器交互的物品。例如，未来的一些冰箱能够与食品上的标签进行交互。这使得冰箱能够扫描你平常需要买的东西，并且当它一旦注意到大量的物品在一定时间内被移走，它将会通知缺少哪些物品，你就可以去购买这些东西。另一个例子是将成为微波的“智能”设备。微波将会从购买到的物品中扫描 RFID 标签，并自动地设置定时器来纠正所需时间。从这些例子中，可以看到杀死标签不是处理消费者隐私的合适的方法。

5.2.2 密码协议

由于智能设备的发明和其他需要再次重复使用标签信息的设备，研究者开始提出一些概念，这些概念可以确保一个人的隐私，在扫描时不禁用标签的情况下，阅读器和标签之间彼此进行认证。其中的一种方法就是借用公钥加密方法的概念来迎合这个准则[8-11]。加密方法是一个简单的，较好定义的算法。接下来将会描述这个类型协议的一个例子。考虑一对匹配的阅读器 R 和标签 T 的公钥和私钥。每个标签首先会嵌入它的阅读器的公钥 R_{pu}，和它唯一的私钥 T_{pr}。在阅读器查询时，标签和阅读器会使用标签的嵌入的密钥进行相互认证。标签 T 将会在一个阅读器公钥 R_{pu}的额外加密内用它的私钥加密一个随机的 $R_{pu}(T_{pr}(n))$，并把它发送给阅读器。一旦阅读器用它的私钥 T_{pr}解码了外部的加密，它将会搜索它的后台数据库，获得标签的匹配公钥 T_{pu}，从标签中解码内部的锁定来检索随机数。阅读器随后会用自己的私钥再次加密这个随机数，与标签相似，它随后用标签的公钥 $T_{pu}(R_{pr}(n))$建立一个在先前周围的一个二次加密回合。在发布这个加密过程之前，阅读器会使用它的密钥，并与当前的标签结合，来形成一个临时的钥匙 T_k，这个钥匙将与随机数一起发送给标签。在检索和解密从阅读器发送来的信息之后，标签将使用 T_k 在它和阅读器之间发送进一步的信息，例如一个人的职位，以便于这些人员访问大楼内特定的位置。图 5-1 描述了这样一个协议。对于基于时间的协议，从标签来的任

何重要信息将会嵌入到第一个发送的消息中，例如供应链系统中产品的EPC号[2]。这个方法阻止了从标签发送信息时不必要的侦听，除非黑客想要检索这个标签的私钥，而这样的检索在短时间内是不可能的。

虽然这个方法的安全性较高，公钥加密（尤其是在RFID标签内执行的公钥加密）需要较强的计算能力，以便于在传输信息的时候进行加密。这不仅增加了标签的尺寸，也显著地增加了每个标签的成本，使这个方法只能在RFID安全的非常有限的领域内使用。

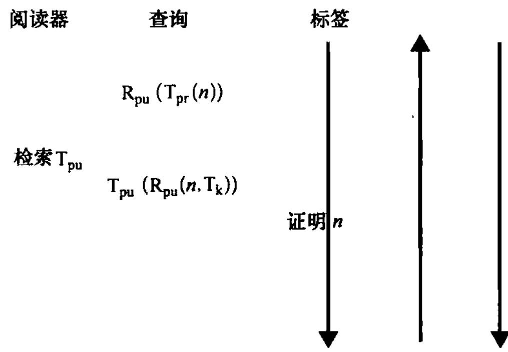


图5-1 加密图协议的一个例子

5.3 RFID隐私保护设备

正如先前说明的，一个RFID标签加长的广播范围，容易被黑客或是任何恶意的第三方利用，进行不期望的读标签操作。当处理隐私保护问题时，这个事实是非常普遍的。这个信息可能是消费者的隐私（消费者先前购买物品的相关信息），或者是一些更需要关注的安全种类（来自一个识别证件中的检索工作或者是政府信息）。为了解决这个问题，先前的许多尝试建议通过一个外部设备来避免对标签不必要的访问。接下来的三个部分将提供这些方法的几个例子。

5.3.1 法拉第笼

处理消费者隐私的方法之一就是所谓的法拉第笼[1]。法拉第笼是一个用铁网或是金属箔片制作的容器，用来阻止一定频率的无线电波。最近，美国政府表示，美国的护照外壳将包含金属材料，来限制射频的穿透，从而阻止封闭护照的长范围扫描[5,6]。尽管如此，这个方法也存在着缺点。主要的缺点是，法拉第笼不能设计成例如手表，容器以及如电视机或是计算机这样较大的物品。这个缺点限制了此方法的使用，限制了例如供应链市场这样的商业投资。

5.3.2　有源干扰设备

另一个用于消费者隐私保护的方法称为有源干扰方法[1]。这个方法允许个人携带一个设备阻止附近的 RFID 阅读器发送或者广播它的信号。但是，如果广播信号的能量过高的话，这个方法是不合法的。这将会使干扰机干扰周围合法的 RFID 阅读器，扰乱了公司的业务。因此，由于这个方法的法律限制，它不太适合作为 RFID 隐私保护的解决方案。

5.3.3　拦截器标签

在过去几年，关于 RFID 安全的问题，已经开展了多方面的研究，如果介绍全部这些研究，将会超越本章的范围。但是，Juels 已经在参考文献［1］中介绍了许多这些技术的细节，以及每个方法的利弊。特别的，将介绍 Juels 技术中的一个，即我们其中一个协议的灵感来源：来自参考文献［12］的隐私位概念。

Juels 在参考文献［12］中使用的方法，与先前描述的干扰方法相似。但是，它的影响不会像运行时那么大，巧妙地与 RFID 单一化协议交互，来扰乱仅仅一些特定的操作。RFID 标签的单一化协议是基于树的方法，阅读器可以在同一时间内扫描来区别多个标签。这个过程是通过重复地查询区域范围内出现的所有标签来工作的，通过保存阅读器接收的一定数量的碰撞，来区分每一个标签。图 5-2 的例子给出了对于这个过程更多的细节描述。图 5-2 为一个三次的树，包含了 8（2^3）个标签序列号，代表了这个树的每个枝叶。假设需要区别标签“001”和“011”。阅读器首先查询所有前缀为“0”的标签。由于两个标签都有这个前缀，所以这个查询将会返回一个碰撞。随后，以“1”为前缀进行查询，此时没有返回任何信号。因此，阅读器不会继续以这个前缀来查询任何标签。单一化协议和之前一样跟随这棵树的碰撞点递归，并最终到达相应的树叶“001”和“011”，对它们的存在，只进行一次响应。

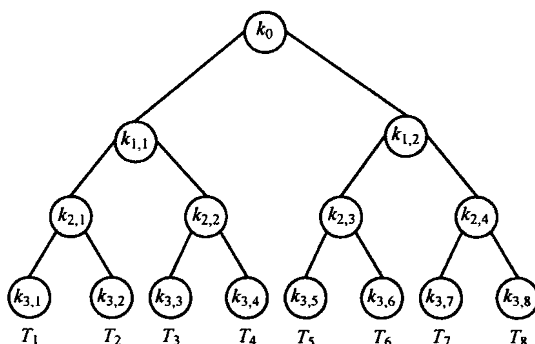


图 5-2　一个分叉树

为了与这个协议进行交互，Juels在参考文献［12］中在标签内使用了一个被称作“隐私保护位”的特殊的位，这个位的值可以是1或0，在这个标签的一个唯一引脚认证之后，阅读器可以容易地对这个位进行转换。当在存储器内，标签存储的值为0，这意味着它有公共的访问，检查标签时，并通过把隐私保护位翻转成1，使标签移动到一个“隐私保护”区。但是只做这些不能保证标签的安全性。一个额外的被称为拦截器的专用标签与标签一起出现，以保护此标签[12]。拦截器标签扰乱恶意阅读器，使阅读器认为出现的标签存在着所有可能的值。标签通过发送阅读器在当前单一化协议内的查询的相应值来具体实现这个功能。例如，如果阅读器要求以“11”开始的序列号，拦截器标签将发送这个值来与实际上可能包含这个值的任何标签发生碰撞，扰乱阅读器到达树的枝叶。这个“扰乱动作”能够使阅读器满负荷（全拦截器），或者是以一种较友好的方式（部分或是选择性拦截器）[13]。但是，友好的拦截器将只会和在一个“隐私保护区内”的树叶进行交互。例如，如先前描述的，标签的隐私保护位在标签认证时，将会从0变成1。友好的拦截器随后只会保护前缀为“1”的标签，如果有必要的话，也允许在这区域内的其他物品得到扫描。在两种方法下，只要拦截器标签出现，RFID标签是安全的。

这个概念得到了具体地设计以用来提高消费者的隐私安全。但是，如果给定了每个标签在信号强度上的差异，那么在不同的零售环境下，它们可以被放置在更好的位置。一个普通的拦截器标签可以在手机中使用，来扰乱试图攻击或者获得相关信息的恶意传输。一个友好的阅读器更适合于供应链市场，可以嵌入到购物袋中。这个安全的临时方法有一个优点，它可以在没有额外处理的情况下，允许先前提及的“智能”设备访问相关的物品。但是，与先前的两个设备相似，这个概念也存在着它的缺陷。甚至一个放置较好的拦截器标签也有失败的可能性，这是由于RFID标签的不可靠传输造成的[1]。而且，阅读器可能最终会利用拦截器标签的弱点，并超过它们信号的强度[14]。为了完全了解关于这个方法的攻击和防御，在考虑任何部署之前，研究和评估仍将持续。

5.4　基于hash函数的RFID协议

由于被设计用来提供对于RFID标签进行有效认证的分离机制的失败，研究者开始开展提供例如在标签自身内的安全方案。一个在许多协议中流行的，被称为hash函数的方法，使用类似加密的方法来实现这样的安全，保护在标签和阅读器之间秘密传输的信息。hash函数是任何数学的函数或者是较好定义的方法，重新排列任何给定的数据成一个合理的小整数，一般为一个阵列提供一个索引来使用[15]。例如SHA-1[16]和MD5[17]这样的hash函数算法，作为在有限的计算范围内传输数据的安全的保护，被广泛地接受。本节探索一个原始的基于hash函数的方法，它是根据在数据库内标签搜索时间的改进而开发的。

5.4.1 hash 锁：原始的基于 hash 函数的方法

协议的其中之一被称为 hash 锁，后来的许多函数都是由它发展形成的[18]。这个基于 hash 函数的方法通过基于 hash 的结果来解锁标签，并得到标签内部的信息。标签由一个“锁定”状态开始，此时阅读器将发送一个锁定值给标签，lock = hash (key)，其中 key 为一个随机值。这个值存储在标签的保留内存位置内（比如一个 Meta-ID 值），标签会进入锁定状态，除了在认证过程中需要的信息，不允许任何其他的信息。为了解锁标签，阅读器必须发送用来形成 Meta-ID 值的原始密钥给标签。在收到这个值之后，标签运行一个 hash 函数，并与它的 Meta-ID 值进行比较。如果匹配，标签将被解锁，在进一步的查询循环中，允许它的 EPC 号[2]响应给阅读器。在对标签内的 EPC 号进行保护时，这个协议非常简单和直观。由于认证过的阅读器只知道每个标签的原始密钥，所以只有它们才能够解锁标签，使这些标签信息进行传输，也能在阅读了码之后，重新锁定标签。

虽然它较简单，这个协议在安全性上存在着一个大的漏洞。因为它需要提供不公开的密钥值，所以它不能在阅读器和标签之间提供相互的认证，而只能够认证阅读器。这个漏洞很容易被恶意的第三方利用来扫描标签，获得它的 Meta-ID 值。这个返回的值会随机地广播到附近的阅读器，并最终会返回特定标签的密钥值。第三方可能随后会使用这个信息来发送给原始标签，获得它的 EPC 号和其他敏感信息。

为了解决上述提及的不足，参考文献［18］的同一作者提出了新的方法，来处理这个随机化。这里需要强调区分每个查询带有随机数的 Meta-ID 值，这样标签和它的值不会容易地被跟踪。因此为了这个方法，一个额外的伪随机数产生器被嵌入到了标签内，以使它位于锁定状态内。但是，这种方法不会存储这个 ID 的 hash 结果，不像每个查询，标签将会用伪随机数产生器给定的随机值来 hash 它的 ID，结果为 hash (ID_k, r)，其中 k 代表在系统一定数量标签中的第 k 个标签，即 ID_1、ID_2、…、ID_k、…、ID_n。

当阅读器查询一个标签时，阅读器将获得两个值。这些值包含标签产生的随机数和 hash 与标签 ID 值比较产生的 hash 函数结果值，以及当前的随机数。为了解锁标签，阅读器必须发送标签原始的 ID 值。因此，阅读器将开始搜索包含所有标签 ID 值的后台数据库，阅读器必须重复地执行一个 hash 函数，以区别标签给定的随机数与每个分离的值。这将允许阅读器的每个 hash 函数结果与标签发送来的一个 hash 结果进行比较。当它们匹配时，阅读器将获得匹配的第 k 个 ID 值，并把这个值发送给标签，使标签解锁。一旦标签进入到解锁状态，任何阅读器可以执行查询来获得标签的 EPC 信息。

除了成功地实现 RFID 标签上信息的安全之外，这个方法也能提供位置隐私的保护。在先前提出的协议 hash 锁定中，每个标签仍然会暴露它的 Meta-ID。但是，这个方法只会指出一个随机号，以及基于这个随机号的 hash 后的值。因此，恶意

的第三方阅读器不能根据标签的Meta-ID来跟踪一个特定的标签（比如商店的一个产品）。在这种情况下，随机hash锁定协议能够提供位置隐私的保护。

与原始hash锁定方案相比，这个随即协议已经有较大的改善，但是这个方法并不能适合所有的情况。由于阅读器必须搜索尽可能多的ID值来寻找匹配的hash结果，当这个系统中有n个标签时，它的运行时间为$O(n)$。因此，当一个系统中标签的数量非常多时，这个方法便存在着扩展性的问题。标签制作伪随机数产生器的额外成本也是这个规模的系统存在的另一个问题。

5.4.2 基于树的方法

正如前面描述的，参考文献［18］报道的基于随机hash函数的方法，在一个系统需要标签迅速增加时，它不能保持少于优化运行的时间和低成本的标签。但是，随着这类系统的流行，所需要的时间将会持续地扩大，需要寻找一种有效的算法来解决规模扩大的问题。在保持它的安全性的同时，尽力地减少基于hash函数的运行时间，参考文献［19-21］的作者提出了一种方法，它能够从线性的复杂度到对数的复杂度，来提高密钥搜索效率。实现上述过程的关键是基于后台数据库的结构，这个结构树是基于树的图设置的。每个节点将需要$O(\log_2 N)$的搜索时间，这已经被多个算法理论所验证。首先使用基于树的方法的小组之一是Molnar等人[21]，他们在这个方法中使用了一个挑战响应协议。这个过程需要多个回合来识别一个标签，每个回合由双方的三个消息组成。但是，由于后台数据库基于树的本质，造成了每个回合需要$O(\log_2 N)$的运行时间，使得每个消息之间形成了较大的通信开销。因此，参考文献［19］改善了这个算法，它缩短了每个回合的长度，使来自标签的消息只有一个，也不提供标签和阅读器之间进一步的交互。为了进一步解释基于树协议的概念，下面将介绍参考文献［19］所实现过程的细节。

一个基于树方法中阅读器的后台数据库包含了用一个二进制方式形成的一个密钥集合。可以参考图5-2中的图。每个节点包含了一个唯一的密钥，每个标签代表一个叶子节点。因此，存在唯一的一条从根到叶子节点的密钥路径，而这些密钥集会被分配到每个叶子节点，这个唯一的路径用来进行认证。例如，标签T_8包含钥匙k_0、$k_{1,2}$、$k_{2,4}$、$k_{3,8}$。当阅读器R认证T_8时，它首先会发送一个随机的n给标签。标签T_8随后会通过执行一个hash来用n加密它的密钥集，然后会把结果发送给阅读器。阅读器随后会通过重复hash发送的随机数n与它的密钥的二进制数，搜索相当于标签的叶子节点，并把这些结果与来自阅读器的加密结果的发送单进行匹配。如果到达叶子节点的路径存在，T_8将被识别，阅读器R会认为这个标签有效。

从上述过程，我们可以得出，到达最终的标签的路径，将会与其他不同的标签的相似的路径共享一定的密钥。例如，标签T_8和T_7共享了钥匙$k_{2,4}$，当然，所有标签将会共享根密钥k_0。使用这种静态结构的一个优势是它的运行时间是对数级

的。例如在图 5-3 中，对一个标签的所有识别只需要 $\log_2(8)=3$ 个搜索阶段。但是这个结构也存在着安全漏洞，即如果一个标签被攻击者控制了，攻击者将获得从根节点到达一个叶子节点的多条路径，包括在这些路径上的密钥。由于这个结构上的密钥从不更新，获得的密钥将仍旧被没有被控制的标签使用，可能会进一步地获得没有被控制的标签的密钥组合。

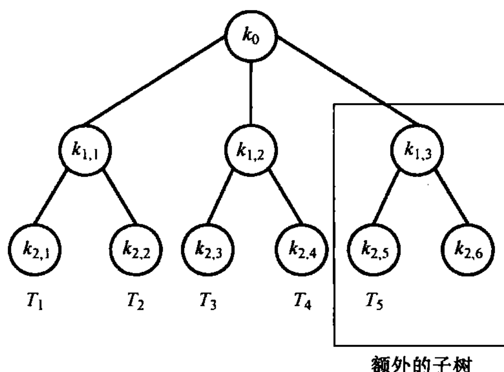


图 5-3　一个有五个标签的静态二叉树

解决上述提及的问题的一个实际方案是在每次认证之后更新密钥。但是，上述提及的静态树结构由于自身的复杂性，不能处理这样的任务。例如，如果需要更新标签 T_1，那么不得不部分或者是全部地改变钥匙 k_0、$k_{1,1}$、$k_{2,1}$ 和 $k_{3,1}$。这将导致独立地影响树的每条路径，不允许其他未认证的标签进行自我更新来使用一个阅读者进行正确的认证。因此，使用这个结构，每个标签需要与阅读器使用的静态树一起，周期性地或是同时得到更新来保持这个方法的同步性。不幸的是，在数以百计或是数以百万计标签的大规模系统中，这个方案并不实际。一个替代的方案是周期性地采集被其他标签每次认证的路径所影响的标签；但是，在试着收集只被一个标签改变它的密钥而影响的大量标签时，这个方法比第一个方法更复杂麻烦。因此，本文作者提出了称为“hash 树”的协议，这个协议是使用于 RFID 系统内隐私认证的一种动态的密钥更新算法。在 5.4.3 节将介绍这个协议。

5.4.3　hash 树：一种动态的密钥更新方法

虽然基于树的方法具有一个高效的搜索时间，但是它们缺乏一个长期的安全保证。由于一个基于树的方法的基础结构，两个或者更多标签之间的路径将共享密钥集。因此，如果一个标签被攻占了，这将导致系统中关于其他标签的信息泄露。为了解决这个问题，阅读器和标签的密钥必须同时得到更新。据我们所知，参考文献［22］是第一篇报道这个漏洞的文献，并对这个问题提出了动态密钥更新方法。在参考文献［22］中，树的后台数据库包含了与参考文献［19］提出的二进制树结构相似的一个密钥集 k。但是，为了更新树，每个节点需要包含一个附加的临时的

密钥 tk 。一开始，每个临时钥匙等于每个节点的当前钥匙的值，即 $tk = k$ 。标签的认证过程实质上是与普通的基于树的方法相同，这个方法中，树会被重复地搜索，来寻找到达代表标签的匹配叶节点的一条路径。在标签认证时，阅读器在 k 上执行一个 hash 函数，并设置它的值为它当前的钥匙值，$k = h(k)$ 。临时密钥随后将一直被更新到事先设置好的当前的密钥值。这样可以确保搜索到达它的叶子节点的路径能够被正确地计算，因为这个协议将在标签的认证期间同时检查一个节点的当前密钥值 k 和临时值 tk 。虽然这个协议能够防止参考文献［19］出现的许多问题，但它只适合于从不会对同一标签扫描两次的系统。这是由于临时密钥的值一次只能提取一个密钥值。如果一个标签被多次扫描，一个节点的临时值 tk 将只提取 i 次扫描一个节点的第 $i-1$ 次。这不仅会允许共享标签路径节点的路径被错误地取消授权，而且最终也会对重复扫描的标签取消授权。因此，本章作者提出了“hash 树”协议，它为基于树的方法提供了一种简单的动态密钥更新系统。这个方法在每次认证时，同时更新阅读器和标签的密钥，也允许先前没有被查询的标签继续得到认证。

hash 树协议由三部分组成：系统初始化，标签识别和系统维护。前面两个部分与参考文献［19］提出的基于静态树的方法非常相似，执行基本的识别功能。但是，不像参考文献［22］描述的动态密钥方法，hash 树协议能够在不更新阅读器的后台数据库的情况下，保护 RFID 系统免受攻击。最后，第三部分用来指导标签进入和退出系统。

hash 协议通过提供参考文献［19］中阅读器和标签相似的范围来进行开始操作。阅读器的后台数据库包含了参考文献［19］和参考文献［22］提供的相似的二进制树形结构。假设 RFID 系统内存在着 N 个标签 T_i（其中 $1 \leqslant i \leqslant N$）和一个阅读器 R，阅读器 R 将分配 N 个标签到平衡二进制树 S 的 N 个叶节点中。树 S 中的每个节点分配一个预先设置的密钥 k 。在引入标签 T_i 时，阅读器将分配 k 个密钥的路径给在树中表示一个没有分配叶子节点的标签。但是，不像先前的协议，这些密钥不会在标签内以它们自然的状态存储。取而代之的是，通过阅读器从原始给定的密钥集得到的每个密钥 k 随机产生的随机数 n 形成一个 hash 结果的阵列，将会与产生那些结果的 n 一起存储到标签 T_i 内。以这种方式存储密钥值对于算法的逻辑来说，是非常重要的，接下来将会简要地介绍这种方式。

正如先前介绍的，这种标签的认证过程与参考文献［19］提出的协议非常相似。但是，关键的不同在于，每个标签是怎样拥有一个预先计算的 hash 结果集，取代了存储在阅读器后台数据库中树的准确密钥。对于来自阅读器的查询，标签将会发送它 hash 结果的阵列，并包括先前描述的用来计算 hash 结果的随机数 n 。阅读器随后会用一个重复的逻辑方式，使用 n 值与从根节点开始的每个节点的 k 值进行计算 hash，直到阅读器到达相当于那个标签的匹配的叶子节点。假设节点的一个路径包含了发送的结果列阵的一个匹配的 hash 结果，这个路径被阻断并发现，随

后标签将会被认证。在标签认证之后，每个标签将会被更新，以阻止先前描述的受控攻击。为了实现这个目标，阅读器将会产生一个不同于标签给定的另外一个随机数 n_2 。这个随机数随后将会被阅读器依照它的密钥树记录的标签 T_i 密钥来执行另一个 hash 函数功能集。对于发送的一个清晰描述，可参照图 5-3 中的标签 T_3 。使用随机值 n_2 ，对于 T_3 的原始标签密钥集的一个连续 hash 函数结果，将会是 $(h(k_0,n_2),h(k_{1,1},n_2),h(k_{2,2},n_2),h(k_{3,3},n_2))$ 这个新的 hash 函数结果清单和随机数 n_2 将随后被发送给标签，来更新它的密钥值。

在解释这一节的系统维护部分之前，我们将解释只更新标签的值而不是阅读器的理论。参考文献 [22] 中提及的密钥，是假设攻击者可以访问标签的秘密值，因此控制了阅读器的后台数据库密钥树系统。这是由于每个标签的密钥集与另一个标签的钥匙集存在一个直接的相互关系，正如先前在 5.4.2 节中，参考文献 [19] 描述的静态树。本章作者提出的算法，将以多种方式来处理此类型的受控攻击。首先，通过更新标签集的值，在一个安全的区域内，它不能再跟踪其他的标签。例如，如果将这个协议用于零售商店的供应链管理，所有客户购买的标签在商店内将不能跟着物品。更重要的是，初始化分配标签和认证之后（比如使用 hash 结果取代原始的密钥数据）的这个过程能够防止攻击者知道连接到阅读器的终端数据库的密钥树中保存的真实密钥值。因为这样，一个标签到另一个标签的密钥位置（甚至是共享一个原始的密钥路径）将不再相互有关联。这就是为什么我们不需要更新阅读器密钥树的主要原因。为了对此进行进一步的解释，在图 5-3 中考虑 T_1 和 T_2 。这两个标签在树结构中共享密钥 $k_{2,1}$ 。在一个静态树中，攻占其中一个标签将获取到其他标签的信息，这是因为标签的每个密钥有它普遍的形式。但是，本章作者提出的算法产生对于这两个标签这个值的不同 hash 结果，因为每个标签将拥有一个不同的随机数（比如 $h(n_1,k_{2,1}) \neq h(n_2,k_{2,1})$ ）因此，攻击者不能够把这些值联系到一起。

最后，在这个协议真实地执行时，用户可能需要同时移除和增加额外的标签到密钥树。因此，将提供一系列的小步骤来执行这个任务。假设一个额外的标签 T_i 需要加入到系统。这项任务可以用两种方式中的一种来执行。阅读器将首先通过叶子节点清单来搜索，如果存在一个有效的空的空间，标签信息将会放置在这个空间处。但是，如果所有的叶节点当前都指向了一个标签，将会生成一个新的子树。根据当前后台数据库中存在的树的深度确定这个子树的长度 d 。这个子树的每个节点将会分配到一个随机产生的密钥，如先前协议对系统的初始化部分描述的，标签 T_i 将用一个随机数分配这些值。图 5-4 给出了当不能够放置一个空节点的附加处理。

因此，用户也同样可以从系统中移除标签。这个任务比较简单，它通过清空标签 T_i 对应的叶节点来完成。如先前图描述的，这允许一个额外的标签可以与那个节点相关。

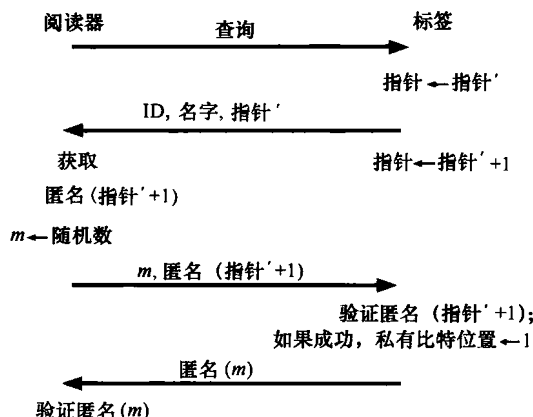


图5-4　一个附加子树的例子

一个安全的认证协议应当符合如下的安全需求[19]：隐私保护，不被跟踪，阻止复制，前向安全，以及拒绝受控。特别是，不被跟踪的需求必须注意的是，标签的输出应当能够与标签的本身相关，否则，这个标签就可能会被攻击者跟踪。纵观本章作者提出的协议，有些人会争辩：尽管每个标签包含着一个不同的随机数 n，多个查询的输出将会是相同的。但是，在认证时每个标签值的更新，将会阻止攻击的发生。正如先前所述，例如，对于一个系统来说，这个协议已经大致能够与这样一个身份识别卡标签交互。因此，假设这个标记每天用来访问一座大楼，对于攻击者来说，跟踪这种类型的标签是非常困难的，除非拥有这个标签的人被跟踪了一整天。另一个是隐私保护问题。对于攻击者获得了标签的当前值和执行一个重放攻击，并不存在一个概要的保护措施。如果对于系统来说它作为一个问题出现，那么每个标签可以简化使用一个5.3节介绍的法拉第笼来保护它，直到此标签被使用为止，这与许多新的电子护照提供的用来进行附加保护的外壳相似[5,6]。通过这样的修改，此协议成功地实现了标签的每个安全需求，因此，使得此协议足够低廉和安全地在实时的RFDI被动标签认证系统中实现。

5.5　其他的RFID认证和隐私保护协议

从先前的方法中发展出了许多的协议，例如大部分基于树的方法（包括我们的“hash树”协议），其他研究者决定采用另一种方法。接下来提及的协议是由Juels[23]提出的，涉及被动标签像有源标签一样，对重新标记的能力。这个协议的思想与Juels先前提出的软-阻止[13]协议一起，分别放在本章的第一部分和最后一部分，“RFID保护：为被动RFID标签设计的认证和隐私保护协议。”

5.5.1 极简的加密

许多主动标签具有采用一种方式来重新标记自己的能力，这使得它们对于第三方恶意攻击者来说是不可区分的，但是仍然能够被授权阅读器认证。由于被动标签没有足够的能量来实现这个功能，许多研究者开始研究使用额外的物品来阻止或是变形来自黑客标签的数据传送，如 5.3 节描述。但是，Juels 提出了称为“极简”系统的一个协议[23]，允许在有限计算能力下，对被动标签进行重新标记。在这样的系统中，每个标签包含了一个小的匿名集；在查询时，一个不同的匿名随后将被给予阅读器，因为每个标签将在每次扫描时翻转它的匿名清单。安全保障在于只有那些已经认证的阅读器才会包含完整的匿名集。一个未经认证的阅读器并不包含一个标签的每个匿名，因此不能够获取任何来自给定的不同表现形式下的标签的安全信息。下面将提供一个此方案的更加整体的例子。

正如先前论述的，每个标签包含了一个匿名阵列 α_i ，其中 i 表示一个标签 m 个匿名中当前的匿名，$1 \leqslant i \leqslant m$ 。但是，单靠这些匿名，这个协议并不安全。如果是这种情况，那么极简系统将容易受到复制攻击，这个攻击是很多为 RFID 标签设计的静态协议容易遭受到的[24]。在这种情况下，攻击者将会查询标签，获得它的当前匿名 α_i ，并重放这个值给阅读器，允许自身被识别为当前扫描的标签。为了解决这个问题，标签只能在阅读器向标签证实了自身以后才能向阅读器进行自身的认证。为了完成这个过程，对于每个匿名 α_i 来说，每个标签包含了两个额外的密钥值 β_i 和 γ_i ，标签和阅读器将使用这些值来认证它们自己。对于阅读器的查询，标签将会响应它当前的匿名 α_i 值。假设这个值是有效的，阅读器将会从后台数据库中找到标签对应的 β_i 和 γ_i ，并发送标签 β_i。对于阅读器的认证，标签将会响应它对应的 γ_i 值。正如你所见的，这个协议模仿了一个简单的挑战-响应协议，但是这个协议是设计用来匿名翻转的。

为了成功地实现标签的长期安全，必须持续的更新标签的 α_i、β_i 和 γ_i 值。在逻辑上，这个更新方法需要在标签和阅读器的每个互相认证之后发生，以便于保持低成本，而不用周期性地一次更新大量的标签。但是，更新标签也存在着新的问题：攻击者可能仍然偷听或者干预着更新过程。为了解决这个问题，Juels 提出了一次一密，这种方法已经用于多个认证协议中，更新一个标签和一个阅读器的当前值。使用这个方法，仅仅进行周期性偷听的恶意第三方将不能获得更新后的 α_i、β_i 和 γ_i 值。

一次一密[25]被认为是比用于密码学中的加密技术更简单的加密方式。因此，它可以使用在比如被动标签这样具有更小计算能力的标签中。一次一密本质上是一个长度为 l 的随机比特串。如果双方共享一个秘密的一次一密 δ，将被证明信息 M 可能在双方之间发送秘密的密码文本 $M \oplus \delta$ ，其中 $\oplus$ 表示异或操作。因此，在标签和阅读器的相互认证之后，阅读器使用这些一次一密来更新 α_i、β_i 和 γ_i 值，并

发送这些密钥给标签，以便于标签同样能更新它的值。假使恶意第三方不能偷听阅读器发送的消息和获得这些密钥，他们不能够获得最近更新的标签值的信息。但是，如果第三方获得了其中一个密钥，这个方案也有一个对于普通一次一密的额外翻转。这涉及通过多个认证会话来使用一次加密。为了实现这个，来自两个不同认证会话的密钥是与更新它的给定标签值 w 异或的，其中 $w \in \alpha_i \cup \beta_i \cup \gamma_i$ 。因此，即使第三方成功地获得了先前会话使用的一个密钥，它也仍然不能获得 w 更新值的任何信息。

极简方法能阻止对公司商业的侦听，例如暗中对供应链市场产品货物的扫描。自从我们第一个提出的在零售环境使用的协议，我们借用了匿名翻转方案的思想以及 Juels 的软-阻止技术[13]，提出了称为“RFID 保护”的协议。

5.5.2 RFID 保护：为被动 RFID 标签设计的认证和隐私保护协议

像沃尔玛、Prcter&Gamble 和美国国防部这样大型的组织，由于它们对于 RFID 作为它们供应链的自动识别工具的需要，最近几年，对 RFID 技术产生了极大的关注[1]。这是因为 RFID 标签与条形码相比具有更显著的优点，RFID 标签能够唯一地识别一种产品类型的单独项，而条形码只能识别每个产品类型。但是，RFID 标签扩大了读范围，使得第三方用户能够偷听在认证的标签和阅读器之间传输的信息，因此存在了一个安全漏洞。供应链管理中的 RFID 安全是被动标签研究和产生不同协议的原始目的，产生了本文作者先前提出的协议：“RFID 保护：为被动 RFID 标签设计的认证和隐私保护协议”。

RFID 保护的思想是由本章先前介绍的由 Juels 提出的两个协议——极简系统[23]和软-阻止技术[13]构成的。极简方法建议每个标签包含一个匿名阵列，在来自阅读器的查询时使用，标签将会翻转它的匿名列表，阻止任何恶意的第三方，因此这些第三方没有完整的列表。为了实现在阅读器和标签之间的相互认证，我们的协议使用了这种匿名翻转概念；但是，不像极简方法，它是不需要更新标签的匿名或者包含额外的密钥来执行这个任务的。原始的软-阻止方法通过使用隐私保护位阻止部分树，与用于变形同时扫描到的多 RFID 标签的分离协议进行交互。我们的协议使用一个隐私位来对 RFID 标签进行隐私保护，代替对单一化协议的交互，隐私保护位在锁定与非锁定状态之间触发标签。在锁定状态期间，标签将只能提供足够的信息来解锁一个认证的标签阅读器，然后此标签会把它的 EPC[2] 发给阅读器，而一个未认证的阅读器将不会包含足够的信息来取回秘密数据。本节剩下的部分提供了我们协议的更多的细节，以及它的安全测量。

RFID 保护协议由四个更小的协议构成：入库协议、登记协议、出库协议和返回协议。每个协议表示了标签预先假定的位置，因此为标签提供了不同等级的安全保护。每个标签包含了多个不同的项：一个匿名清单，一个代表标签当前指向的匿名的指针，一个代表一个产品名的号码，作为第一位包含在标签 EPC 中的隐私保

护位，以及在认证时包括前面提到的 EPC（在这个协议中指 ID），一个标签需要发送到阅读器的所有秘密信息。假设直到每个标签到达零售环境时，都不存在进一步的交互，因此标签将以入库和登记协议开始。这通过一个标签的隐私保护位来表示，这个隐私保护位以 0 开始表示标签处于非锁定状态。入库协议是登记协议中的前两个步骤（而在出库协议中是返回协议的前两个步骤），因此我们只介绍这两个协议。

在登记协议中对一个标签的查询时，标签复制它当前的指针，并发送这个指针、标签的名称和 ID 号给标签，在发送完信息后，增加它的指针额外的时间。指针在每个查询时是如何增加的，将在接下来的协议中解释。正如先前论述的，假设标签在到达商店之前，没有进一步的交互。因此，在仍然解锁状态时，任何超过那个点的查询，被假设为是一个认证的阅读器给定的，所以在给出它的 ID 号之前，不需要额外的安全保护。在接收到标签的信息后，阅读器将使用标签的类名，并在给定指针的位置后提取匿名，匿名［指针 + 1］。只有经过认证的阅读器才会知道在给定指针后返回被给的匿名，而不是现在的匿名。因此，标签的认证之后，标签将会翻转它的隐私位到 1，表示标签处于“锁定”状态，标签将会进入返回协议，直到它再次进入非锁定状态。但是，这个过程发生之前，阅读器也要发送一个额外的随机数 m ，来表示在标签内一个指针的位置，其中 $1 \leqslant m \leqslant n$ ，n 为一个标签内的匿名数。在阅读器确认之后，标签也将发送匿名［m］，这个匿名是根据标签的位置来定的。这将能够使协议克服标签复制问题。一些复制的标签可能包含一个认证标签的有效匿名，但是不包含全部的这些匿名，或者没有把这些匿名以正确的顺序放置。如果这是真实的，随机匿名的查询将有很大的机会抓住这点，因此如果匿名［m］不是有效的，或者标签没有在设计的时间周期内发送一个匿名，阅读器将会报警，表示客户持有非法的或者复制的标签。图 5-5 中给出了这个标签的一个示例图。

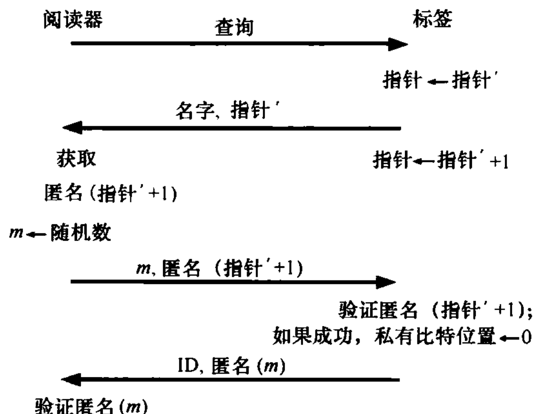


图 5-5　RFID 防范的登记协议

正如先前论述的，在标签完成了登记协议之后，它将通过设置隐私位为1来进入锁定状态，随后它将进入返回协议过程，直到标签被解锁为止。在这个锁定状态，假设标签出现在了恶意的第三方阅读器内，将不能像先前协议那样较为容易地发送信息。因此在查询时，标签将用指针重复同样的过程，但是只包含了它的副本、名称，不包含ID。指针在每次查询是如何增加的，在这个协议中比先前的协议更加普遍。任何试图重复扫描标签信息的黑客将会触发标签持续地开关它当前的指针。因此，如果黑客在登记协议中提取了来自阅读器的一个或是两个匿名，它将会扰乱黑客，使黑客不知道这些匿名的位置。这将使阅读器从偷走的匿名和没有确认的标签中，更好地发现复制的标签。回到返回协议的问题，有效的阅读器将会获得从标签发来的信息，并且重复发送标签一个随机指针 m 和匿名［指针+1］来确认标签自身的过程。在阅读器确认后，标签将不止发送匿名［m］，也会同样发送它的ID。这里发送匿名［m］的目的不是来检查标签的复制，而是检查不道德的标签。一个有效的阅读器只能提取对应这个协议的商店内的标签，如果客户正放回一个物品，可能会破坏标签，来增加它原始值的价值，使客户付出到更多的现金，或者将他的物品换为更高的价格。因此，这种随机匿名特性在客户能够从商业中偷走钱之前，试图发现这些败坏的标签。图5-6给出了这种标签的示例图。

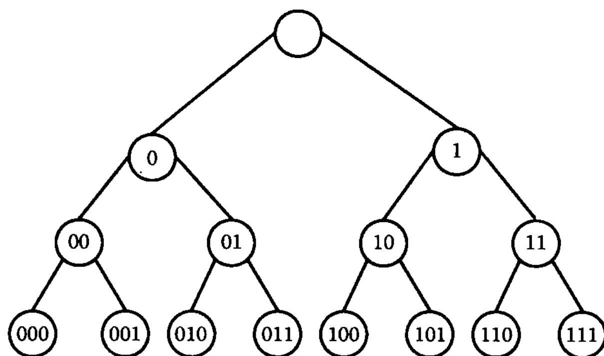


图5-6 RFID防范的返回协议

这个协议实现了在参考文献［19］前列出的许多安全需求。然而，为何标签内的特定项在协议的特定部分会泄露，这还不是非常清晰的。通过查看登记协议，有人会怀疑：为什么在阅读器生效之前ID会泄露，或者为什么在两个子协议上给定了多个匿名。这两个问题的原因是，我们对这两个协议（至少在认证阅读器状态下）作了一个假设，即它们在一幢大楼内执行。因此，给定了被动RFID标签的有限广播范围，协议执行一个扩展的时间周期，人们很难偷听到传输的信息。此外，因为我们的协议包含在一个物理建筑物中，我们假设了零售商店具有一些安全保护机制，以阻止未经认证的阅读器进入商店。这可以通过在靠近入口处安装监测设备来检测未经认证的阅读器来实现[26]。

另一个在这个协议中需要讨论的安全问题是避免跟踪攻击问题。正如在讨论 hash 树协议时论述的，跟踪涉及需要持续地接收来自标签的一个静态值，来得知这个标签的行踪。这个问题处理用于搜索协议中的标签信息的类属名。解决这个问题不需要额外的设备方法，但是需要理解类属名值代表什么。这个名称代表了产品类型，而不是在商店内的每个产品。例如，两瓶苏打水只要它们是同一个公司生产的，可以有相同的类属名。如果协议不以这样的方式设置，系统的后台数据库使用永远不需要返回的额外的时隙，如每瓶苏打水的信息，这将会在物理上和内存上潜在地增加成本，需要使用额外的时隙返回这样单独每瓶苏打水的信息。

在上述相关的安全需求的解释上，RFID 保护具有成为了进一步研究的潜在标准零售环境协议的潜力。

5.6 结论

RFID 是一种非常有前景的技术，它能够改变我们的生活方式。但是，在它成为现实之前，必须考虑像消费者隐私保护，欺诈保护以及监测这样的安全问题。本章只介绍了 RFID 技术保护领域当前安全所使用的很小的一部分协议。在 RFID 的使用正式超过条形码之前，必须继续进行大量的研究，以保证对单个标签有效的识别和保护。但是，仍然需要相信的是，不是每个使用被动标签的系统会采用相同的方法，但是在一定程度上，许多安全协议通过每个分离 RFID 系统的 EPC 准则，将成为标准。在 RFID 安全领域存在着大量的余地来改善和创新，使 RFID 技术融入到我们每天的生活。

参考文献

1. A. Juels, RFID security and privacy: A research survey, *IEEE Journals on Selected Areas in Communications*, 24(2): 381–394, 2006.
2. EPCglobal. Epcglobal website. http://www.EPCglobalinc.org/, 2007.
3. S. Stern, Security trumps privacy, Christian Science Monitor, 2001.
4. D. Molnar and D. Wagner, Privacy and security in library RFID: Issues, practices, and architectures, in B. Pfitzmann and P. McDaniel, eds., *Proc. ACM Conf. Commun. Comput. Security*, pp. 210–219, Washington, DC, 2004.
5. International Civil Aviation Organization ICAO, Document 9303, Machine readable travel documents (MRTD), Part I, Machine readable passports, 2005.
6. A. Juels, D. Molnar, and D. Wagner, Security and privacy issues in e-passports, in *Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm)*, pp. 74–88, Athens, Greece, September 2005.
7. G. Barber, E. Tsibertzopoulos, and H. B.A., An analysis of using epcglobal class-1 generation-2 RFID technology for wireless asset management, in *Military Communications Conference*, vol. 1, pp. 245–251, Atlantic City, NJ, October 2005.

8. S.E. Sarma, S.A. Weis, and D.W. Engels, RFID systems and security and privacy implications, in *Workshop on Cryptographic Hardware and Embedded Systems (CHES) 2002, LNCS no. 2523*, pp. 454–469, Redwood Shores, CA, 2003.
9. M. Ohkubo, K. Suzuki, and S. Kinoshita, Cryptographic approach to privacy-friendly tags, in *RFID Privacy Workshop*, MIT, Cambridge, MA, November 2003.
10. S. Bono, M. Green, A. Stubblefield, A. Juels, A. Rubin, and M. Szydlo, Security analysis of a cryptographically-enabled RFID device, in *USENIX Security Symposium*, pp. 1–16, Baltimore, MD, July–August 2005, USENIX.
11. J. Wolkerstorfer, Is elliptic-curve cryptography suitable to secure RFID tags?, in *Hand-out of the Ecrypt Workshop on RFID and Lightweight Crypto*, Graz, Austria, July 2005.
12. A. Juels, R.L. Rivest, and M. Szydlo, The blocker tag: Selective blocking of RFID tags for consumer privacy, in *Proceedings of the 10th ACM Conference on Computer and Communication Security*, pp. 103–111, Washington, DC, 2003.
13. A. Juels and J. Brainard, Soft blocking: Flexible blocker tags on the cheap, *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society*, pp. 1–7, Washington, DC, 2004.
14. M. Rieback, B. Crispo, and A. Tanenbaum, RFID Guardian: A battery-powered mobile device for RFID privacy management, in C. Boyd and J. M. Gonzlez Nieto, eds., *Proceedings of the Australasian Conference on Information Security and Privacy*, Springer-Verlag, New York, 2005, vol. 3574, *Lecture Notes in Computer Science*, Brisbane, Australia, pp. 184–194.
15. I. Mironov, Hash functions: Theory, attacks, and applications, Microsoft Research, Silicon Valley Campus, November 2005.
16. National Institute of Standards and Technology, Secure hash standard, Federal Information Processing Standards Publications (FIPS PUBS), April 1995.
17. R. Rivest, The MD5 message-digest algorithm, MIT Laboratory for Computer Science and RSA Data Security, Inc., April 1992.
18. D. Johnson, C. Perkins, and J. Arkko, Mobility support in IPv6, RFC 3775, IETF, June 2004.
19. T. Dimitriou, A secure and efficient rfid protocol that could make big brother (partially) obsolete, in *PERCOM 06: Proceedings of the Fourth Annual IEEE International Conference on Pervasive Computing and Communications*, pp. 269–275, Washington, DC, 2006. IEEE Computer Society.
20. D. Molnar, A. Soppera, and D. Wagner, A scalable, delegatable pseudonym protocol enabling ownership transfer of rfid tags, Cryptology ePrint Archive, Report 2005/315, 2005. http://eprint.iacr.org/.
21. D. Molnar and D. Wagner, Privacy and security in library rfid: issues, practices, and architectures, in *CCS 04: Proceedings of the 11th ACM Conference on Computer and Communications Security*, pp. 210–219, New York, 2004. ACM.
22. L. Lu, J. Han, L. Hu, Y. Liu, and L.M. Ni, Dynamic key-updating: Privacy-preserving authentication for RFID systems, in *PERCOM 07: Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications*, pp. 13–22, Washington, DC, 2007. IEEE Computer Society.
23. A. Juels, Minimalist cryptography for low-cost rfid tags, in *Proceedings of the 4th International Conference on Security in Communication Networks*, vol. 3352, pp. 149–164, 2004.

24. S.E. Sarma, Towards the five-cent tag, Technical Report MIT-AUTOID-WH-006, Auto-ID Labs, 2001. http://www.autoidlabs.org/.
25. A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1996.
26. T. Li and R. Deng, Vulnerability analysis of emap-an efficient RFID mutual authentication protocol, in *International Conference on Availability, Reliability and Security*, Vienna, Austria, 2007.

第 6 章　RFID 的安全问题

如今普适计算正在快速地发展，在普适计算中，具有较少计算资源的设备越来越重要。其中的一部分设备是各种传感器，最值得注意的是无线射频识别（RFID）标签。这种快速增长的状况激发了在这种环境下使能和维持安全的轻量级协议的需求，而这些轻量级协议并不是一项简单的任务。因此，本章将提供对于 RFID 安全领域的概述，从基础定义和相关的场景开始，以及对 RFID 安全协议数量评估的适当机制。接下来，这个领域的主要方案将和它们的缺点将会被一起给出（其中的一些协议的缺点到现在还没有被指出）。基于这些，介绍新的非决定性（ND）的加密协议，它设计用来提供 RFID 环境下的安全。最后，给出这个领域的展望，包括对开放问题以及将来期望的趋势的描述。

6.1　概述

当前的计算趋势已经转到了能够实现普适计算模型的无线通信上。在这样的环境中，大量的设备将只具有有限的计算资源、处理能力以及存储或是能量供应。普适计算中主要的代表是 RFID 标签。

RFID 设备预计不久将会应用到普适计算环境的众多通信设备中，这些应用领域包括零售业和健康保障系统等。因此安全保护将变得日益重要，这不仅是从用户的观点和期望出发，也是从法律的角度出发而考虑的。因为 RFID 设备只有较少的资源，这就意味着基于 RFID 计算环境的安全保障是一个显著的挑战——存在着设备方案的许多严格的安全需求，而这些设备是缺乏处理、存储和通信能力的。

本章将针对 RFID 安全问题介绍一个扩展的研究。在 6.2 节，将给出基本的定义和适当的参考场景。在 6.3 节中，介绍领域的当前状态，其中描述具有已知弱点的现存的协议，还有一些新出现的不足问题。进一步的，本节介绍用于 RFID 系统的这些协议的量化评价标准。依据这几小节，6.4 节将介绍两个新的协议。这些新的轻量级的协议是非确定性的（ND），并且很好地适应了 RFID 环境的规定（这些规定的简要分析，也同样在本节给出）。6.5 节给出对预期未来的 RFID 系统中的研究场景的展望，并在 6.6 节给出总结。在本章的最后，列出相关的参考文献。

6.2　基本定义和参考场景

RFID 由 RFID 标签作为主要部件的前端，包括阅读器和连接数据库通信链路

的后台部分，以及数据库。分界线代表了一个阅读器，假设它属于系统的后台部分。此外，经常假设后台部分是安全的，前台部分将从安全的角度来分析，本章中也是这样假设的。图 6-1 所示为这个环境的情况。

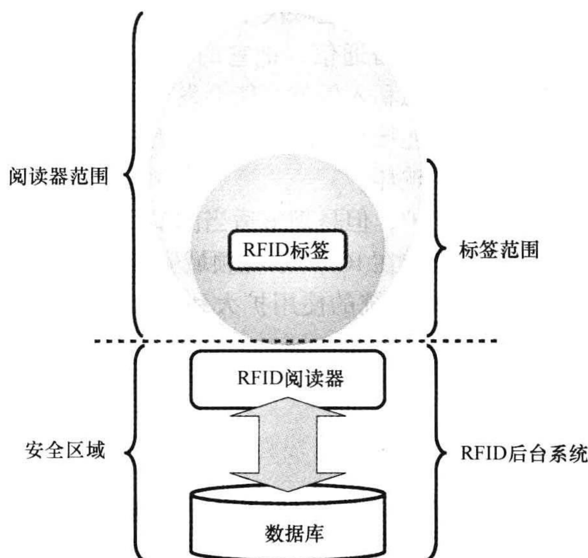


图 6-1　RFID 参照环境场景

RFID 标签由识别号（ID）编码的一个微芯片和一根天线组成。标签和阅读器之间的通信通过电磁耦合在射频频率发生。阅读器在标签电路内部引起电压，为标签提供充足的能量来执行必要的计算和响应。这种功能特性的标签称为被动标签。但是，标签也能够具有由电池提供的自主能量，这些标签则称为有源标签。被动标签较为便宜，它们的操作距离只能达到 3m，而且具有一个相对高的错误速率。另外，有源标签价格较高，但是操作距离可以达到数百米，并具有较低的错误速率。两种标签均可只读，或者一读多写，或只写。由于成本的限制，市场上的大部分标签是被动标签，本章也将集中介绍被动标签。

现在根据参考文献 [4]，安全保护意味着使设备和资源受到的安全漏洞降到最小。这是通过包括原始密码、逻辑和物理机制等安全机制的部署来实现的，以达到以下安全服务的实现。

1）认证，确保对等通信的实体是声明的实体。

2）机密性，即阻止数据非认证泄漏。

3）完整性，即确保任何对数据的修改都将会被检测到。

4）访问控制，即阻止未经认证使用资源。

5）不可抵赖性，对拒绝承认自己制造信息内容的信息发起者提供其发起信息的证明。

6）登记和审核，即能够对可疑活动进行检测，并对成功的入侵进行分析。

关于上述的定义，已经给出了额外的说明。在被动标签中，认证常常是被一些外部的资源触发的。这种类型的认证被称为强迫认证。强迫认证是非常重要的，因为它可能导致隐私威胁——当然，这也取决于环境。例如，只要物品在商店的货架上，标签可以自由地与周围的设备通信，把它的ID发给周围的设备。但是，当用户购买了贴有标签的物品时，隐私入侵就可能会发生，解决这个问题的办法是，把用户的标识与收银机处的标签连接在一起。然后，无论何时未授权的阅读器发生了强迫认证，隐私保护都将会被破坏。

虽然本章并不涉及隐私保护，但是理解适当的安全设备是需要的，所以上述的解释是非常有必要的，因为当前在RFID系统领域安全措施的使用主要目的都是在于对隐私的保护。但是，RFID系统的使用扩大和新安全问题的出现，也会在本章其余的部分进行讨论。

6.3 领域的当前状态

本节首先给出原始密码的相关问题（即安全机制）的概述，以及加密协议相关问题（安全设备）和安全度量问题。

6.3.1 原始密码问题概述

安全实现的主要障碍是成本。最近的RFID实现期望具有以下的特性[5]：它是被动电源供应的，具有存储标签ID号的96位只读存储器，这个ID号对于每个标签来说，都是独一无二的。芯片每秒能进行200次的读操作。据估计，在经济上可以接受的范围内，能够分配到最大2000个门操作来确保安全。考虑到摩尔定律（作为一个比特传统），现在上限已经可以达到4000~5000个门操作。

这就加大了对安全协议的需求，而这些安全协议也必须是轻量级的。虽然许多文献中报道的协议都声称是轻量级的，但是它们是基于许多潜在的假设的，假设实现不需要额外的门操作。例如，它们假设单向的hash函数自动地具有轻量级协议的资格。但是像MD-x或者SHA-x类这样的大多数协议，并没有轻量级的特性[6]。此外，协议中每个额外的步骤，常常需要额外的特定电路；安全协议内的步骤是语义相关的，结果是，每个额外的步骤导致了在RFID标签上需要实现更加复杂的算法。

因此只能选择特定的隐藏原语，包括大多数明显的是轻量级的AES[4]和轻量级的DES（DESL）[7]。在轻量级AES中，大致相当于需要3400个门操作，电路被优化来进行低功耗的操作。作者声明，在DESL中实现了与AES可以相当的能力，少了45%的芯片尺寸，降低了86%的时钟循环以及大致1800个门操作的使用。后者可以在144个时钟循环内加密64位的明码文本。DESL特别适合我们的目标。它将

是产生 128 位长的 hash 值的基础（参考文献［8］中报道的对于单向 hash 函数所使用的对称时钟加密的各种原理）。

6.3.2　密码协议问题概述

由于上述限制，强调认证的安全协议需要包含尽可能少的步骤，其中一个简单的“请求-响应”就有一个最理想的结构。如果在一个协议中需要更多的回合，那么额外的消息在句法上相等会更受欢迎。这意味着在不同的输入下，可以使用相同的电路。这也与串行化计算要优于并行化计算是一致的。被动标签最重要的是每个时钟循环的能量消耗，这意味着需要能耗最小化。因此并行计算应该由串行计算所替代[9]。

在 RFID 协议领域能够识别下列常见的威胁。

1）被动攻击可以通过简单地监视阅读器和标签之间的通信来获得足够的信息。

2）中间人攻击是攻击者用自己的数据修改请求来实现的。一个恰当选择的请求可能会误导标签或是阅读器去相信它们之间是直接通信的，而在实际中不是这样的。而消息被攻击者移交和修改了。

3）主动攻击是攻击者主动地卷入到通信中，并修改消息（中间人攻击便是一种动态攻击）。

4）重放攻击是攻击者记录交换的信息，在没有必要知道信息中包含内容或是如何计算这些信息内容的情况下，简单地重新使用这些信息。

5）重放攻击是将攻击者信息中继部署在标签和阅读器之间，进行信息中继，它使得阅读器错误地相信标签与它相距很近，因此它能够采取相应的行动。

6）恶意阅读器攻击有许多种类，但是我们将关注于未经认证的跟踪攻击，其中恶意系统在没有必要知道标签真实身份的情况下，跟踪这个标签，但是仅仅是基于其对请求响应的情况下，在各种地方识别标签。

7）物理攻击是从电路中直接读出标签的内容，而不是使用无线连接。

6.3.3　RFID 安全的一些重要的密码协议

RFID 加密协议可以分为单个标签协议和多个标签协议。此外，这些协议中的每个组能够分为单回合协议和多回合协议。接下来将介绍从参考文献［10］中总结的多回合单标签协议。

（1）Weis、Sarma、Rivest 和 Engels 协议[11]　一个阅读器和一个标签共享一个密钥 x。被阅读器触发后，标签产生一个随机数 r，并计算字符串 $(r,(ID \| H(ID) \oplus f_x(r)))$，这个字符串将会发送给阅读器（这里，“$\|$”表示字符串的互相关联，“$\oplus$”表示异或操作，$f_x$ 是使用 x 作为参数的一个伪随机函数）。在确认之后，阅读器回复标签的 ID。很显然，在第三步暴露 ID 号，这是明显有疑问的，更不用提及，重放攻击是没有意义的，因为第一个和第二个消息加密图是独立的。

（2）Henrici和Muller协议[12]　在一个标签计算了$H(ID)$和$H(s○ID)$之后，阅读器发送一个请求。接下来，它将发送$H(ID)$、$H(s○ID)$和δ_s给阅读器（其中“○”表示一些选择运算符，“s”是步骤的数目，“δ_s”是当前的和先前的会话数的差，当先前的执行有效时，这个差等于1），在接收到这个消息之后，阅读器为标签计算一个新的ID($ID \leftarrow ID○r$)，并更新数据库，发送r和$H(r○s○ID)$给标签。在收到之后，标签能够确认r的完整性，并能够计算新的$ID(ID \leftarrow ID○r)$。因此，标签和阅读器处于同步状态下。但是，这个协议的一个问题是，如果异或用来进行“○”操作时，攻击者能够实现阅读器解同步。在这种情况下，攻击者通过一个0位字符串在第三步中替换r，结果是，获得了$H(r \oplus s \oplus ID) = H(s \oplus ID)$。这个值与来自第二步的值相同，可以发送给阅读器，也能够被攻击者读取。因此，当标签检查来自第三步的信息时，标签将用一个与阅读器计算的不同的值来更新它的新ID。结果是数据库去同步化。

（3）Ohkubo、Suzuki和Kinoshita协议[13]　这个协议中，标签和阅读器共享两个hash函数G和H，以及一个初始化密钥s_i。阅读器发送一个请求给标签，这触发了标签通过计算$H^1(s_i) = H(s_i)$来计算一个新的密钥，并存储这个新值。同时，标签计算$G^1(s_i) = G(s_i)$，并发送这个值给阅读器。后台数据库hash每一个存储的秘密值，找到一个匹配对$(ID, G^1(s_i))$。在第二次运行中，$H^2(s_i) = H(H(s_i))$，并且$G^2(s_i) = G(G(s_i))$被计算和使用。但是，这个版本易受重放攻击，因为标签发送给阅读器的第二个消息没有链接到第一个消息。因此，攻击者能够能够发送一个请求给标签，并在随后的时间记录使用它来响应给阅读器的信息。Avoine、Dysli和Oechslin提出了一个解决方案，方案中第一个消息包含了一个新的请求r，第二个消息作为$G(s_i \oplus r)$被计算[14]。

（4）Molnar和Wagner协议[15]　这是预计没有任何缺点的协议的一个例子。标签和阅读器共享一个密钥x。首先，阅读器选择随机数r_r，并发送它给标签。标签选择随机数r_t，计算$\sigma_1 = ID \oplus f_x(0, r_r, r_t)$，并把它发送给阅读器。阅读器通过计算$ID = \sigma_1 \oplus f_x(0, r_r, r_t)$来使用$\sigma_1$提取ID，然后用$\sigma_2 = ID \oplus f_x(1, r_r, r_t)$来回复。在接收到这个信息之后，标签检查ID是否是好的。

但是，即使是后两个协议，也仍然存在着一些先前未描述过的不足。

1）关于修改的Ohkubo、Suzuki和Kinoshita协议，如果一个攻击者总是发送相同的请求，来自标签的响应将总是相同的。这对于强迫认证来说，是一个严重的问题。因为尽管不知道标签的标识，但它仍然可以被跟踪。

2）Molnar和Wagner协议的改进版本的缺点与第三个消息相关。如果标签在接收到这个消息后没有匹配的，那会发生什么？如果期望标签回应，这显然意味着需要一个具有额外步骤的更复杂的协议，但是没有这些步骤。没有了这些额外的步骤，使用参考文献［10］中给出的协议，这意味着攻击者能够调整第二个消息中的一些位，阅读器将会确定一个错误的ID。因此，协议不能确保已检查ID

的完整性。

正如参考文献［10］中提到的，很多实施异或函数的协议可能会通过提交一个作为计算的输入的零向量（所有的位是0）被成功袭击。原因很清楚——通过与零向量异或的所有位序列生产一个相同的位序列。在这种情况下，检查输入是否不是零向量是明智的。当一个单位向量（所有的位为1）用来异或，导致输入的位序列无效时，原因相似。

上面描述的协议是确定性的协议。Hoper 和 Blum 采用一种不同的方法利用 HB 协议（它的后续协议是 HB +[17] 和 HB ++ 协议[18]）[16]。所有的 HB 变种都是基于伴有噪声问题的奇偶校验位学习（LPN）。这个问题在给定多个 $b_i = a_i * x \oplus v_i$ 的计算结果之后，这里 v_i（也称为噪声）等于1，它的值也有可能在区间$[0,1/2)$之中，要求攻击者计算一个在阅读器和标签之间共享的 k 位的秘密 x。由于噪声为 1 的可能性严格小于0.5，所以攻击者能够用一些选择的 a 连续多次地来请求一个标签。一旦获得了具有线性独立的 $a - s$ 的 k 等式，x 能够通过高斯消元法来进行恢复。此外，采用这个原理是以主动攻击为基础的，而这对于 HB 协议类是没有抵抗力的。像 HB ++ 中间人这一类协议有一些其他的不足，在参考文献［19］中描述了这些不足。

接下来是一些其他的具有代表性的单标签协议，参考文献［10］给出了它们缺点描述的扩展概述。另外，参考文献［10］也讲到了多标签协议，多标签协议是用在一个阅读器领域内需要同时出现两个标签这样的场景中的。

6.3.4 测量密码图协议的轻量级特性

基于迄今给出的事实，在 RFID 环境中明显存在着与安全相关的一个重要问题。实际上，假设所有上面描述的协议是轻量级的，因为只有这样的协议，才适合在 RFID 系统中的实现。我们将开放性地讨论对于轻量级协议需要哪些特性。为了提供这些特性的量化评价标准，参考文献［20］提出了适当的方法论。接下来也将简要地描述它的背景。

计算机科学中最广泛使用的理论模型之一就是图灵机。它的目的在于研究什么是理论上可以计算的。另外，这个模型能够进行计算复杂度（关于时间和空间）的测量，将其定义为一个增长性的输入函数。但是当设计加密图协议时，将面临着工程上的问题，上面的模型结果是不合适的。

第一，密码图协议的情况中，输入的尺寸通常是较小的，并且与渐近性的行为无关。

第二，需要考虑一个具体的真实的计算架构，因为新设计的协议将在这个架构上运行，因此，它的架构需要反映出当前计算设备的总体特性。

第三，它不寻址通信过程，但这对于 RFID 系统是非常有必要的。

此外，当前的计算结构仍然被冯·诺依曼模型很好地描述。但是，把 RFID 系统用这种模型来看待是不合适的。第一，这些设备只包含较弱的处理器，即处理部

分不包含经常支持一般处理器的功能。第二，这个架构不能直接映射到轻量级协议的问题领域。第三，通信同样没有被很好地覆盖到。

RFID设备由两个基本的电路组成。中心部分是带有存储能力的处理器部分（这个部分包含了逻辑门），而第二个部分是无线射频通信电路（RF部分）。因此，一个协议轻量级特性的测量需要覆盖以下几点：它应当支持任何的逻辑功能，应当包括通信成本，也应当考虑当前的制造技术。

基于以上的这些需求，图6-2给出了一个模型。考虑到量化，实现特定协议需要的NAND门数量正是为了达到目的。只把NAND门作为量化的基本可能乍一看有点儿奇怪。但是，这是布尔代数基本事实之一，即任何逻辑功能能够用布尔函数的一个逻辑完全集来实现。此外，这反映了技术的实际性，因为大多数逻辑电路都是以这种方式实现的。

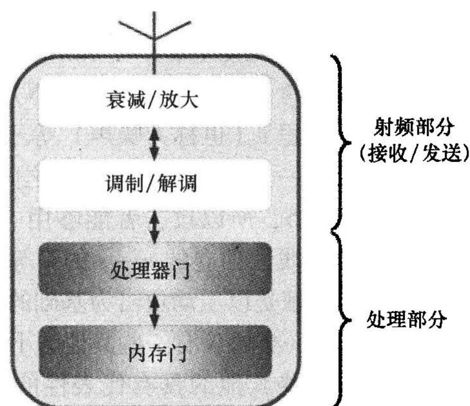


图6-2 为派生轻量级协议标准的RFID模型

布尔函数的一个完全集由与和非组成，并且能够用与非门来实现。这个实现需要一个与非门来得到“非”，以及两个与非门来得到“与”（与非布尔函数也是在布尔代数上为逻辑完全的）。最后，应当考虑RFID需要的时钟成分。这些事实代表了轻量级协议量化的基础。

进一步的，每个协议的实现需要一些存储单元，为了这个目的，D存储单元被选择来量化，因为它们是时钟定时的。每个D单元，即翻转，需要5个与非门[21]。除了存储外，对于安全需要（密码图协议）的典型逻辑函数是按位异或的，并且加法模2^n。变量x、y按位异或可以获得$(x\bar{y})$和$(\bar{x}y)$。在一些优化步骤后，得到4个与非门。同样的，一位全加器需要11个与非门来实现。对于执行模2^n的加法操作，需要使用n倍的11个与非门。这个解释对于本章的目的来说是足够的，更多用与非门来实现布尔函数的细节，可以在参考文献［21］中找到。

一个稍微困难的问题是对通信部分量化的引入。通信部分实际上是对消息编码，并通过电磁耦合在一定的距离内收发消息码。协议越复杂，将要传输的位的数量也就越大。这些位可以被认为存储在一些存储器中，并利用这个存储器，在发送器和接收器之间传输。因此，可以引入和与非门一起实现的D存储单元用于量化通信。

总结以上内容，量化射频部分将变成需要传输的位的数量，并且这些位的代价将会用需要存储它们的D与非储存单元的数量来测量。现在形成一个轻量级协议的定义，是有可能的了。我们引入代价N，它包括实现一定协议需要的与非门数量。这意味着N包括存储S、处理P和通信C门，即$N=S+P+C$。这个消耗N用

来作为对评估轻量级协议的一个量化。

一个协议的代价N是通过实现它而需要的与非门的总数量来测量的，这包括了存储、处理和通信部分。

上述定义对于比较轻量级协议的代价是有用的。为了表示一个协议是否是轻量级的确切界限，是需要一个背景的，并且取决于技术的实际状况。当前，RFID系统环境中轻量级协议的一个合理的界限是消耗2500个与非门。1500个门是用来支持RFID加密操作的，另外1000个门是在通信中用的（值得注意的是，在当前RFID系统中阅读器使用的门总数量大约是5000个）。

6.4 新的非确定性加密图协议

本节引入的协议适合于单标签和多标签应用。它们是非确定性的协议，意思是，当阅读器得到标签的响应时，这个响应期望的值位于一定的间隔内。阅读器需要在这些间隔内检查所有可能的分散的值，并找到一个匹配的。这样的协议把主要的计算任务都放在了阅读器和后台系统上。在大多数情况下，这是可以接受的，特别是在后台具有非常大的计算资源的RFID架构的情况下（这样的原理在现今非常的普遍，特别是在数字签名领域）。

6.4.1 第一个非确定性协议

第一个非确定（ND）协议的运行如图6-3所示。阅读器和标签共享一个密钥 x_i，并且双方都能够计算一样强的单向的hash函数 H。标签和阅读器同样共享确定随机值 Δt 计算间隔的 n（这个 n 不需要保密）。

现在发送认证过程。

1）阅读器用一个时间戳来请求标签。

2）接收到时间戳之后，标签在接收到的值中，选择性地检查它。假设RFID标签具有自主的时间电路，以及为了阻止回复，前面的值需要进行存储，这是不合理的。由于有限的资源，存储接收请求的存储器是FIFO的，例如，是由4个位置组成。当接收第五个值时，第一个值将被覆盖。所以如果时间标是新的，标签存储这个值，并从间隔 $[0, n-1]$ 中计算随机数 Δt，意思是，它可能有 n 个不同的值。标签将密钥 s 与 $(t \oplus \Delta t)$ 连接起来，并hash这个字符串。

3）标签根据先前的步骤发送结果给阅读器。

4）接收到来自第二步骤的信息后，阅读器开始计算以找到一个匹配。为了找到这个匹配，阅读器用来自 (ID,s) 对的s-es来计算 $H(s \parallel (t \oplus 0)), \cdots, H(s \parallel (t \oplus (n-1)))$，这个 (ID,s) 对是存储在数据库中的。如果找到了一个匹配，那么标签将会被认证。

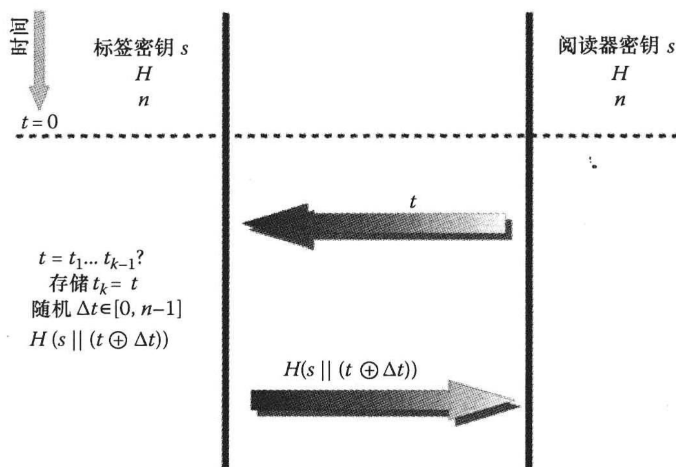


图6-3 第一个ND协议

在第二个步骤中随意地检查一个随机请求的新鲜性来防止强迫认证。如果一个恶意阅读器不断使用相同的请求，标签的响应将会总是相同的，从而这个标签将是可跟踪的。当然，由于标签有限的资源，所有存储请求的清单一般是相对短的，但是需要的存储器在将来将会增长，这个步骤随后变成了义务性的。但是，本节最后的分析显示，对新鲜随机性的检查在可行的技术内已经是强制性的，特别是如果请求分配了48位，这意味着在4个96位的存储器位置中，可以存储8个请求。

6.4.2 第二个非确定性协议

第二个ND协议如图6-4所示。阅读器和标签能够计算一个强的单向hash函数H。标签给定了一个密钥s，而这个s也已经被阅读器所知道。再次，标签和阅读器共享确定随机值Δt计算间隔的n。

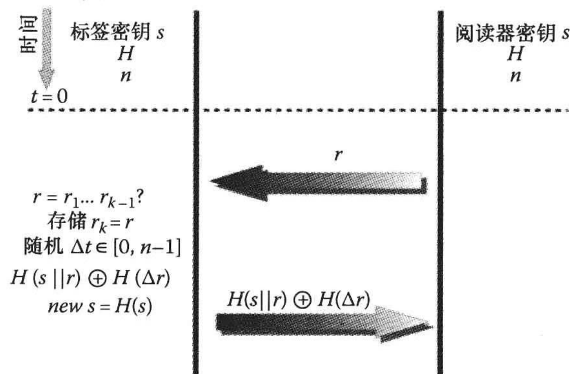


图6-4 第二个ND协议

现在，认证过程如下。

1）阅读器发送一个请求 r 给标签。

2）接收到之后，如果接收的 r 在已经使用的请求清单上，标签则选择性地确认。如果不在，标签会存储接收到的请求，并且计算随机数 Δr。接着，标签计算 $H(s \parallel r)$，并且进一步通过用 $H(\Delta r)$ 异或它来随机化这个结果。标签发送这个结果给阅读器。

3）在接收到来自先前步骤的消息后，阅读器计算 $H(s \parallel r) \oplus H(\Delta r=0),\cdots,H(s \parallel r) \oplus H(\Delta r=n-1)$，直到找到了一个认证标签的匹配。当然，阅读器已经用 (ID,s) 对接入到了一个数据库。

这个协议有一些重要的特性需要讨论。这是基于一个随机数请求，标签选择性地检查请求的新鲜性，以阻止来自恶意阅读器的强制认证。再者，由于在将来可用内存的增加，这个步骤随后会变成必需的。接下去，在请求被检查之后，它将与密钥 s 连接在一起，只在随后进行 hash 操作。把 hash 函数应用到两个变量 s 和 r 常常是通过异或 s 和 r 来进行的（这个例子是对 Okhubo、Suzuki 和 Kinoshita 协议[14]的修改）。如果 r 只有 0 比特位组成，或是全部被设置位 1，这将会造成问题。前面的情况时，异或操作将导致 s 本身取反，而后面的情况时，异或操作将导致对 s 取反。通过系变量来取代对他们进行异或，这个问题将会被避免，并且也不需要额外的电路来检查是否在请求中所有的位都是 0 或 1。最后，这个结果用 hashΔr 来异或。这可以支持多个标签的应用。通过偏移产生 Δr 电路中的初始值，（对于）Δr 的两个值将会区别，这将会是最终的结果。因此攻击者将不知道标签实际上是最后双标签来进行响应。HashΔr 的原因是，这些值的间隔相对较小，例如，被定义成 8 位。没有 hash 的话，只有 $H(s \parallel r)$ 的最后 8 位将会受到影响。

6.4.3　非确定性协议的简要分析

在进行细节介绍之前，需要再次进行论述的是假设阅读器是处于安全的环境中，并且是作为系统后台部分。

非确定性协议的安全：对于 RFID 环境的两个非确定性协议，可以判定的是：所有的消息是唯一的，把随机数看成第三部分，有选择性地检查新鲜以及响应（当然，在工程实际中，这种唯一性会受到限制，这是因为有效的 FIFO 存储器数量和有限的扩展的 Δt 和 Δr 间隔。）。

此外，这两个协议中的第一个消息把密码图绑到了第二个消息。因此，主动和被动攻击会被阻止。回复攻击也同样会被阻止。由于持续地改变消息，恶意阅读器跟踪将是不可能的。但是，阻止物理的攻击仍然是一个未解决的问题[22]。

还有，回复攻击能够通过增加一个由 Hancke 和 Kuhn[23]提出的距离绑定协议来阻止。这是可能的，因为非确定性协议是逻辑独立于目的在于绑定的 Hancke 和 Kuhn 协议的，但是非确定性协议的目的在于对强迫认证的认证和阻止。

资源的消耗：接下来将对实现上述非确定性协议需要的与非门数量进行定量估计。

1）存储1位需要5个门（假设是D触发器的情况）。

2）由于事实上，分组密码能够用来进行密码图hash，并假设实现的是DESL，这需要大约1800个门[7]。虽然hash使用的分组加密对于一般的实现不是高效的，但在RFID系统的情况中，它是合理的（意味着hash函数在软件上实现，将比在一般的计算设备上实现要快，但是如果它们在硬件上实现，它们将显著地超过DESL所需要门的数量）。

3）值Δr和Δs和一个实现一起产生的，是适合轻量级目的的，并且部署了一个转换寄存器[24]。在这个实现中，转换寄存器具有一个异或反馈环，反馈环中，一个输入是转换寄存器的输出，而另外一个输入在寄存器中是第n位。异或门的输出被送入转换寄存器的第一个存储单元。一个m位的寄存器能够标识2^m个不同的值，但是全0的情况在电路中不能出现，因此实际不同值的数目为2^m-1个。通过选择一个适当的n，结果序列是伪随机的，并且具有最大的长度，如果m使得$p(x)=x^m+x^n+1$束缚于GF之上的话。一个4位的转移寄存器需要大约60个门，而8位则需要大约120个门，以此类推。因此假设我们的实现需要8位，就需要120个门。

4）随机的，当$k=4$时，需要n位位置来存储使用的r和t值，以及1位位置来存储保密的S（ID)。因此，假设96位的值，这意味着需要$(4+1)\times 96=480$ bit和2400个门。注意的是，一个n位位置能够存储3个请求r和t，如果这些是32位长的，那么在这种情况下，存储请求的总数量将是12。

按位异或需要4个门，因此异或128位需要512个门（这是指第二个协议，而第一个协议只需要$8\times 4=32$个门。）

使用上述的值，整个消耗估计需要4800个门，其中不包括需要用来比较接收请求的新鲜的逻辑门。因此主要的消耗在于需要进行新鲜检查的存储单元。如果不包括可选的新鲜检查，门的数量大约为2400个，而这可以使上述协议保持为轻量级。但是，甚至包括可选的步骤，这两个非确定性协议仍然属于RFID的轻量级实现。

为了表明目的，6.3.3节讨论的大多数RFID协议的轻量级是如何计算的。更精确的，让我们集中关注Molnar和Wagner协议，这个协议包含的结构，对于其他三个协议来说也是典型的。因此，在Molnar和Wagner协议的情况中，标签的响应如下：$\sigma_1=ID\oplus f_x(0,r_r,r_t)$。ID的存储需要480个门，$r_r$的存储需要$5\times 8=40$个门，其中假设$r_r$的长度为8位。同样的操作对于密钥$x$也是正确的——如果假设它的长度是8位，需要40个门来存储x。此外，8位长度r_t以及一个转移寄存器的计算需要120个门，使用DESL的f_x的实现需要1800个门。最后，异或96位需要$4\times 96=384$个门，因此这个协议需要门的总量大约是2864个。这也表示了非确定

性协议提供安全设备的效率，包括了对恶意跟踪的阻止。

6.5 RFID 安全的开放性问题

在 2006 年，对 IPTS 进行的研究表明，社会接受和相信 RFID 系统的比例是非常低的，这对于广泛部署 RFID 系统[25]来说，是一个较大的障碍。此外，CapGemini 的一个调查显示，用户发现 RFID 安全是一个问题，因为它们与其他一些如会员卡[26]等隐私保护有效的技术相比，更容易被入侵。因此 RFID 系统的安全问题在最近几年已经受到了较大的关注。根据这个领域最具有代表性的一些网站[27]，在 2002 年这个领域只发表了 1 篇文献，2003 年是 11 篇，而 2007 年是将近 50 篇。

2006 年，Rieback 等人提出，切实解决方案需要的主要安全问题是：标签上的加密，密钥撤销，标准化以及立法[28]。确实的，在过去的两年时间里，正如已经提及的，研究领域已经提出和评估了多个密码的解决方案[27]。但是，许多解决方案仅仅只是理论上的，并没有实际地实现过（这里的原因已经在先前的几节内进行过分析）。另外，甚至那些已经实现的解决方案也并没有在真实环境中测试和评估过[29]。

虽然安全问题以及隐私保护问题已经被工业界所知，但是到目前为止并没有足够的驱动力使得在大量的标签生产中实现安全保护的解决方案。这是很现实的假设，即安全特性在标准中占据了充分的位置后，它们也将会在标签中实现。

一般地，与 RFID 系统相关的安全问题很大程度上取决于使用标签的类型。结果是，标签的性能不同，解决方案便会不同。被动标签只有很小的空间，能够进行对加密解决方案的实现，所以它不能使用传统的安全保护解决方案，以及像前面几节描述的特殊的轻量级解决方案。标签的尺寸和能力越大，标签就可以实现更有效的安全解决方案。另外，更多的能力也带来了新的问题需要解决。在简单的被动标签上，它会将能力集中在保护读过程上。相反的，在具有更大能力的多目的有源标签上，安全测量还不得不考虑到来自读、写和擦操作的接入控制。

接下来的几节，将讨论未来五个最重要领域存在的问题：物理上的 RFID 系统保护，原始密码，标签的密码协议，后台强调的专门应用问题，法律问题，以及一般的开放问题。

6.5.1 RFID 系统的物理安全

如果攻击者在物理上接入标签，RFID 标签则能够被多种方式跟踪。标签的电路可以用尖锐的物品来拆卸，并且连接到的天线也可以割断。（这是作为一个隐私保护测量[30]被 Karjoth 和 Moskowitz 提出的）。此外，标签也能够被一个大的电磁脉冲破坏。由于电磁脉冲没有在物理上跟踪标签，所以不可能注意到标签是否被破坏。当然，标签的使用者希望有一种方法可以让他们较容易地知道标签是否在正常

工作。在阻断机制的帮助下，当标签自身未受损，允许更加灵活的标签使用时，可以阻止读标签[31]。另外，在一次应用中，破坏或是阻止读标签是不可接受的，并且是需要被阻止的。因此，需要适当的标签保护的方法。

标签也能够在物理上与贴着标签的物品分离，这是没有意义的，但是这也是最严重的威胁之一。不久，分离的标签能够被附着到一个新的物品上，因此识别系统会被破坏。这是事实，因为把RFID标签附着到物品上，又阻止它被去除，这是不可能的。再者，对于一些领域，标签能够被去掉是很重要的。而且，当附着标签的物品包含一个有效的标签时能够被识别，这也是很重要的。

对标签物理的访问，使得旁路攻击成为了可能，如时域和能量分析攻击。Oren和Shamir提出了一种能量分析攻击，这个攻击不需要物理上的交流。这个攻击能够在没有攻击者的情况下执行，并且标签不需要传输任何的数据，使得这种攻击很难被检测到[32]。因此，需要解决这类旁路攻击的解决方案。最后，一个著名的，通常是用来防止篡改的方法，它是Ross Anderson在10多年前提出的，这种方法对RFID标签系统也是有效的[20]。

6.5.2 原始密码和加密协议

对于阻止强迫认证的安全协议，已经进行了大量的研究（比如保护标签，以阻止非认证的阅读）。许多已经提出的协议是依靠hash函数来作为一个轻量级的解决方案。Feldhofer和Rechberger指出，这些现存的hash函数对于保护简单的被动标签是不可行的（这是因为这些hash函数需要数千个逻辑门来存储内部的向量）。一种对称的类似加密的DESL能够在许多情况下用来产生hash值。由于一个完全的hash函数具有良好的性能，他能够在一个RFID标签中实现，所以它仍然很受欢迎。

另外一个需要额外研究的重要假设（这个额外的研究假设已经被许多协议开发者解决）是能够使用质量随机数产生器。但是这个领域的研究并没有非常多的进展。来自auto-ID实验室的研究者已经公布了他们的解决方案，这个方案有可能解决这个问题[33]。但是考虑到解决方案已经被公布，使得它是否能够抵抗住其他研究者找到的安全漏洞，成了一个疑问。

当前大多数轻量级的安全保护解决方案，主要关心的是保护标签的身份，而其他的包括数字签名在内的重要问题，还没解决。这主要是由于非对称加密似乎不处于现存的简单被动标签的范围之内，但是由于技术的发展，这个问题将会很快被提上议程。

6.5.3 后台系统

RFID系统的后台包括从RFID阅读器到数据库的信息传输，以及对于这个数据库的使用。后台保护不会存在RFID标签的限制，因此可以使用保护通信和数据

库的传统安全保护措施。读一个标签的可能性是一种相对有限的安全保护威胁，虽然访问包含多个 RFID 标签数据的大数据库常常是有较大的维数（特别是对于隐私保护来说）。

因此，需要仔细地设计后台系统的保护。特别的，阅读器是一个有趣的目标。攻击者可能强迫恶意软件入侵阅读器，它将会把数据泄漏给入侵者，甚至是彻底地破坏后台系统。为了更好地理解这类攻击，需要注意到 RFID 系统的响应可能包含了不是它的 ID 号的内容，而是一些命令。知道了阅读器是直接地连接到数据库，一个简单的 SQL 命令就能够造成各种类型的破坏。例如 SQL 命令“; shutdown-”将会导致服务拒绝攻击。此外，数据库完整性也会被通过一个 SQL 语句删除其相关语句而遭到攻击，例如“; drop table <someAttackTable>”。最后，使用描述的代码注入技术将 RFID 系统和 SQL 世界相连，甚至可以设计出病毒来（当然，这种情况是在 RFID 系统可写的情况下）[34]。

标签的数据与其他数据在数据库内的结合，导致出现了一些新的附加问题。在这被称为面包屑威胁中，用户与存入到数据库的物品的 RFID 标签相关[35]。当这个用户把物品拿走时，数据库内的关系不会改变。因此，以后当 RFID 激活的标签物品被发现时，用户可能被带到了一个错误位置。

由于不同的应用存在不同的需求和不同的安全威胁，我们将注意力转移到对于特定应用的研究。例如，护照使用的标签与使用在产品物流上的标签相比应该具有更高的防篡改能力。最后，阅读器的存储器以及阅读器内可能的临时文件也会成为某些恶意攻击者的目标。

随着 RFID 标签使用的新领域的出现，像需要所有权转移问题的解决方案的新需要也将会增长[36]。

6.5.4 法律问题

法律问题是一个永久性的挑战，并已经在文献的许多地方讨论过（比如，可以在参考文献［37］中看到）。一旦描述 RFID 调节的合法的争论转化为法律，一个安全的解决方案将随之而来。

需要注意的是，像加利福尼亚这样的许多州以及欧盟已经对 RFID 领域的法律问题展开争论。欧盟的情况是，他们已经开始讨论是否确实需要对 RFID 建立专门的法律[38]。虽然已经给出了一个官方的论定，不会颁布 RFID 相关的特定法律（因为法律是技术中立的，并且需要更加含蓄的覆盖技术），事情似乎仍在变化。RFID 系统的安全是如此的重要，并且它是专门的技术，以至于 RFID 不可能在法律上进行彻底地抽象。因此欧盟委员会现在考虑通过提供一些“软性法律导引”来进行一个折中[39]。尽管如此，许多法律已经保留使用到普适计算环境中，其中最受人关注的法律之一就是数据隐私指令[40]。

6.5.5 一般的RFID安全问题

当基本的原始密码和协议足够轻量级时，对安全问题发展的注意力开始转移到了专门化应用的问题上。更复杂的标签将会出现新的安全威胁类型，反过来，这也就需要新的解决方案。这个判断理由的一个例子是对于标签写数据的可能性，这使得影响阅读器的恶意软件的侵入成为可能[41,42]。另外一个预测上述趋势的例子是由 Fishbin 和 Roy 提出的[43]，这个例子中，标签应当提供更多的信息给阅读器，使得这些标签能够比其他没有提供更多信息的标签与阅读器之间更加紧密。这个概念的可行性，以及它如何来抵制攻击者设计的不同种类的专门设备，是公开的。

威胁的阻止也需要适当的威胁模型。Rao 等人提出了一种简单的模型，它能够用来分析威胁的程度[44]。这种威胁评估模型，虽然比较肤浅，但是可以作为进一步研究在不同环境中的更好的威胁分析方法的基础。

除了技术上的，也同样存在着需要解决的非技术上的问题。Garfinkel 等人指出用户应当了解更多的关于 RFID 标签的生产以及标签如何工作的信息[35]。他们建议在不同的情况下定义 RFID 标签行为的公共政策。这些政策将给用户相关的信息，用这些信息可以来评估一个给定标签提供的侵害隐私的危害。Ayode 提出阅读器读标签时，应当发出声音，这与数码相机在拍照时发出声音相似[45]。这样一来，用户可以知道什么时候标签被读取，从而提高了这项技术的可接受性。

最后，继续解决关于 RFID 标签法律相关的问题是非常重要的。在部署这些标签时，很容易会产生带有敏感个人信息的庞大的数据库。RFID 系统可能会导致对这类数据库的额外清理规则的需要，以及受到影响的人员如何被适当的通知到，并且应当阐明相关的责任问题[46]。例如，如果商店没有在客户购买了物品之后毁坏物品内的 RFID 标签，那商店需要负什么样的法律责任？再如果毁坏失败了，这是标签制造商的责任，还是商店的责任？法律问题可能将为工程师带来研究机会，以便于开发在将来提高 RFID 标签安全使用的解决方案。

6.6 结论

由于 RFID 系统使用范围的扩大，RFID 解决方案的安全问题已经变得越来越重要。为了确保在这种环境中的安全，已经探讨了许多相关的问题：基于安全设备部署安全机制的技术问题，用户安全重要性的意识，以及法律问题。

通常情况下，信息系统安全不同领域内的问题首先是以技术的观点进行看待，这在 RFID 系统中也是正确的。因此本章首先从安全技术的观点开始，并对它进行广泛的涵盖。给出了当前大多数典型和重要的协议。这些协议的目的是确保合法的可以接受的认证，以及相关的数据完整性（比如数据库同步）。本文介绍了迄今文献报道的这些协议的不足。另外，也描述了这些协议中一些新发现的不足。

考虑到这点，我们提出了两个新的轻量级加密图协议。这些协议是非确定性的，并且需要最小化标签端的资源。为了实现这个，大量的计算被放在了阅读器，即后台端，这在大多数信息系统中是可以接受的。由于这些基本的特性，这些提出的协议通过阻止强制认证来提供高效的安全保护。我们简要分析了它们对已知攻击的抵抗能力。我们相信这两个协议能够实现它们预期的使用目标。此外，本章介绍了能对轻量级协议提供定量评估的设备。这样的定量评估对于 RFID 环境是必要的，因为这些设备具有非常有限的资源，安全需要考虑非常严格的需求来提供。最后，本章讨论了可能在不久的将来提上议程的一些悬而未决的问题。

参考文献

1. Gartner Group, RFID market $3 billion in 2010, *RFID Update*, Dec. 13, 2005, http://www.rfidupdate.com/articles/index.php?id=1014.
2. G. Roussos, Enabling RFID in retail, *Computer*, 39(3), 25–30, 2006.
3. D. Trček, Security and privacy in RFID based wireless networks, in *Handbook of Research on Wireless Security*, eds. Y. Zhang, J. Zheng, M. Ma, Vol. II, pp. 723–731, IGI Global, New York/Hershey, 2008.
4. International Standards Organization, Information processing systems: Open systems interconnection—Basic reference model, security architecture, part 2, ISO 7498-2, Geneva, 1989.
5. S. A. Weis, Security and privacy in radiofrequency identification devices, unpublished Masters thesis, MIT, Cambridge, MA, 2003.
6. M. Feldhofer and C. Rechberger, A case against currently used hash functions in RFID protocols, in *Workshop on RFID Security Security 06*, Graz, 2006, http://www.iaik.tugraz.at/aboutus/people/feldhofer/papers/RFIDSec06_slides.pdf.
7. A. Poschmann, G. Leander, K. Schramm, and C. Paar, New light-weight crypto algorithms for RFID, in *Proceedings of the IEEE International Symposium on Circuits and Systems-ISCAS 2007*, New Orleans, LA, 2007.
8. B. Schneier, *Applied Cryptography*, 2nd edn., John Wiley & Sons, New York, 1995.
9. M. Feldhofer, J. Wolkerstorfer, and V. Rijmen, AES implementation on a grain of sand, *Information Security, IEE Proceedings*, 152(1), 13–20, 2005.
10. S. Piramuthu, Protocols for RFID tag/reader authentication, *Decision Support Systems*, 43(3), 897–914, 2007.
11. S. A. Weis, S. E. Sarma, R. Rivest, and D. W. Engels, Security and privacy aspects of low-cost radio frequency identification systems, in *Proceedings of the 1st Security in Pervasive Computing, Lecture Notes in Computer Science*, 2802, 201–212, Boppard, Germany, 2004.
12. D. Henrici and P. Muller, Hash-based enhancement of location privacy for radio-frequency identification devices using varying identifiers, in *Proceedings of the 1st International Workshop on Pervasive Computing and Communication Security*, pp. 149–153, Orlando, FL, 2004.
13. M. Ohkubo, K. Suzuki, and S. Kinoshita, A cryptographic approach to a 'privacy-friendly' tags, in *RFID Privacy Workshop, MIT*, November 15, Cambridge, MA, 2003.

14. G. Avoine, F. Dysli, and P. Oechslin, Reducing time complexity in RFID systems, in *Proceedings of the 12th Annual Workshop on Selected Areas in Cryptography*, pp. 291–306, Kingston, Canada, 2005.
15. D. Molnar and D. Wagner, Privacy and security in library RFID: Issues, practices, and architectures, in *Proceedings of the 11th ACM Conference on Computer and Communications Security*, pp. 210–219, ACM Press, Washington, DC, 2004.
16. N. J. Hopper and M. M. Blum, Secure human identification protocols, in *Advances in Cryptology ASIACRYPT 01, Lecture Notes in Computer Science*, 2248, pp. 52–66, Gold Coast, Australia, 2001.
17. A. Juels and S. A. Weis, Authenticating pervasive devices with human protocols, in *Advanced in Cryptology—CRYPTO'05, Lecture Notes in Computer Science*, 3126, 293–308, Santa Barbara, CA, 2005.
18. J. Bringer, H. Chabanne, and E. Dottax, HB++: A lightweight authentication protocol secure against some attacks, in *IEEE International Conference on Pervasive Services, Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing*, Lyon, France, 2006.
19. H. Gilbert, M. Robshaw, and H. Sibert, An active attack against HB+—A provably secure lightweight protocol, *IEE Electronic Letters*, 41(21), 1169–1170, 2005.
20. D. Trček and D. Kovač, Formal apparatus for measurement of lightweight protocols, *Computer Standards and Interfaces*, 31(2), 305–308, 2008, http://dx.doi.org/10.1016/j.csi.2008.02.004.
21. L. Vodovnik and S. Rebersek, Digital circuits, Faculty of Electrical Engineering, Ljubljana, 1986.
22. R. Anderson and M. Kuhn, Tamper resistance—A cautionary note, in *Proceedings of the Second USENIX Workshop on Electronic Commerce*, pp. 1–11, Oakland, CA, 1996.
23. G. P. Hancke and M. G. Kuhn, An RFID distance bounding protocol, in *Proceedings of the IEEE/Create-Net SecureComm*, pp. 67–73, Athens, Greece, 2005.
24. P. Horowitz and W. Hill, *The Art of Electronics*, Cambridge University Press, New York, 1989.
25. M. van Lieshout, L. Grossi, G. Spinelli, S. Helmus, L. Kool, L. Pennings, R. Stap, T. Veugen, B. van der Waaij, and C. Borean, RFID technologies: Engineering issues, Challenges and policy options, in *IPTS*, Sevilla, 2006, http://ftp.jrc.es/eur22770en.pdf.
26. Capgemini, RFID and consumers—What European consumers think about radio frequency identifications and implications for businesses, Capgemini report, 2005, http://www.capgemini.com/news/2005/Capgemini_European_RFID_report.pdf.
27. G. Avoine, RFID security and privacy lounge, UCL, Louvain, 2008, http://www.avoine.net/rfid/.
28. M. Rieback, B. Crispo, and A. Tanenbaum, The evolution of RFID security, *IEEE Pervasive Computing*, 5(1), 62–69, 2006.
29. J. Ayoade, Roadmap to solving security and privacy concerns in RFID systems, *Computer Law and Security Report*, 23(6), 555–561, 2007.
30. G. Karjoth and P. Moskowitz, Disabling RFID tags with visible confirmation: Clipped tags are silenced, in *WPES '05: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, pp. 27–30, Alexandria, VA, 2005.

31. A. Juels, R. L. Rivest, and M. Szydlo, The blocker tag: Selective blocking of RFID tags for consumer privacy, in *Proceedings of the 10th ACM Conference on Computer and Communications Security*, pp. 103–111, Washington, DC, 2003.
32. Y. Oren and A. Shamir, Remote password extraction from RFID tags, *IEEE Transactions on Computers*, 56(9), 1292–1296, September 2007.
33. W. Che, H. Deng, X. Tan, and J. Wang, A random number generator for application in RFID tags, *Networked RFID Systems and Lightweight Cryptography*, pp. 279–288, Springer, 2008.
34. M. R. Rieback, B. Crispo, and A. Tanenbaum, RFID malware: Truth vs. myth, *IEEE Security & Privacy Magazine*, 4(4), 70–72, 2006.
35. S. L. Garfinkel, A. Juels, and R. Pappu, RFID privacy: An overview of problems and proposed solutions, *IEEE Security & Privacy Magazine*, 3(3), 34–43, 2005.
36. K. Osaka, T. Takagi, K. Yamazaki, and O. Takahashi, An efficient and secure RFID security method with ownership transfer, in *International Conference on Computational Intelligence and Security*, pp. 1090–1095, 2006.
37. D. Flint, RFID tags, security and the individual, *Computer Law and Security Report*, 22(2), 165–168, 2006.
38. S. Pritchard, CeBIT 2007: Europe opts out of RFID regulation, PCPro, Dennis Publishing Ltd., March 15, 2007, http://www.pcpro.co.uk/news/107699/cebit-2007-europe-opts-out-of-rfid-regulation.html.
39. e-practice.eu, Commission launches consultation on radio frequency identification (RFID), March 3, 2008, European Communities, http://www.epractice.eu/document/4426.
40. European Commission, Privacy and electronic communications directive, 02/58/EC, *Official Journal of the European Communities*, L201, July 31, 2002, Brussels, 2002.
41. M. R. Rieback, P. N. D. Simpson, B. Crispo, and A. S. Tanenbaum, RFID malware: Design principles and examples, *Pervasive and Mobile Computing*, 2(4), 405–426, November 2006.
42. S. Ortiz Jr., How secure is RFID?, *Computer*, 39(7), 17–19, 2006.
43. K. P. Fishkin and S. Roy, Enhancing RFID privacy via antenna energy analysis, Technical Report Technical Memo IRS-TR-03-012, Intel Research, Seattle, WA, 2003.
44. S. Rao, N. Thantry, and R. Pendse, RFID security threats to consumers: Hype vs. Reality, in *The 41st Annual IEEE International Carnahan Conference on Security Technology*, pp. 59–63, Oct. 8–11, Ottawa, Ontario, Canada, 2007.
45. J. Ayoade, Roadmap to solving security and privacy concerns in RFID systems, *Computer Law and Security Report*, 23(6), 555–561, 2007.
46. E. P. Kelly and G. S. Erickson, RFID tags: Commercial applications vs. privacy rights, *Industrial Management and Data Systems*, 105(6), 703–713, 2005.

第7章 RFID的部署：供应链案例研究

RFID系统的部署需要仔细地分析、计划和控制，以使相关机构获得一个优化的方案，否则，预期的好处可能不能够实现。本章将介绍RFID的基本原理和部署方法，讨论研究一个供应链案例。案例涉及在一个RFID测试中对各种成员中的货盘和纸箱的跟踪和识别。部署方法分为三个方面：商业、基础设施和部署环境。理解影响机构决策行为的当前因素和部署的动机，将在部署方法的第一阶段进行评价。一旦机构核定了一个商业案例，RFID技术需求的投入将在物理环境下开始实施。如果部署的动机仍然是积极的，那么将会进行RFID测试和实验阶段。阶段过渡性的具体因素以及结果的具体化，将会在整个部署方法中进行引导组织。

7.1 概述

RFID的部署需要一个能够成功进行初始化和部署的方法。本章将讨论使用一个全国性的供应链（NSC）案例作为一个例子来讨论逻辑部署的方法。这个方法的不同之处在于，它关注于为什么组织正在考虑和部署RFID技术，以及RFID技术是如何影响是否已经完成了部署的决策。这种方法是比较灵活的，一个小的企业或是一个完整的供应链都能够实现这个方法。值得注意的是，这个方法关注的是来自RFID结构的数据的采集，并没有把数据集成到企业的信息系统中去。

与其他商业处理和IT基础设施的改变相似，RFID部署需要一个商业承诺，以及一个完整的分析和计划。整个过程需要编写一些相关的或是组织自身和RFID顾问提出的文档。

一个组织可能存在于一个环境中，这个环境决定RFID技术的实现。另外，组织或是企业可能会初始化这种类型的识别系统的检查。组织的无条件行为将取决于激励环境。如果技术已经处于统治地位，那么无条件行为的范围将会受到限制。例如在英国，Marks&Spencer已经托管了RFID的实现，但是也同样提供RFID设备[1]。在本章的案例研究中，参加者希望形成一个团体，来研究在供应链中RFID数据交互的可能性。这个依据是通过给参与者商业利益，在一个系列站点范围间，用已经识别的物品的增长的实时知识数据研究电子产品代码（EPC）网络。这个范围包括了跟踪物品，显示所有权的变动以及物品在商业点之间的移动。

本章讨论了RFID基本依据和部署方法的三个阶段。在7.2节的阶段一：商业环境中，检查了组织前景的，这个前景有关于商业案例中产生RFID部署的阶段转换动机（PTM）。在7.3节阶段二：基础环境中，调查了RFID的物理和技术需要，

这必须是与商业案例一起理解的，以便于强化或是发展一个使用 RFID 的过程。在 7.4 节的阶段三：部署环境中，讨论了原型测试和试验阶段，最后在 7.5 节给出了结论。工作案例研究方法中有关的每个过程将被持续解释。整章中，数据采集需求，系统集成或是系统部署，以及 RFID 部署的可能商业加强将通过等级来予以评定。

7.2 第一阶段：商业环境

商业环境检查组织的当前状态，以及如何通过激励环境修正它们的无条件行为。如果 RFID 是被托管的，那么组织有义务采用新的技术，并把这个功能集成到所需的商业案例中。存在这样一些环境，这些环境中托管会与组织的生存冲突。因此，企业可能不会去处理托管，不得不去寻求其他的供应和或是用户。因此在这类情况中，组织的成本将会比优势大。如果这个组织正在分析用 RFID 可以对其加强的非统治性商业案例，若代价超过了收益，那它可能会结束它的研究。在先前的这两个例子中，对组织来说，PTM 的参数对于执行方法的下一个阶段不是令人满意的。

图 7-1 所示，为 RFID 依据和部署方法第一阶段的概述。这个阶段检查了激励环境，分析了组织考虑部署 RFID 技术的原因。激励环境将会评估商业案例是否是一个来自强制的指引。组织将会显示出激励环境特性导致的不同行为的变化的度。利用来自这个阶段的 PTM 的结果来测量引发进入下一个阶段。

7.2.1 商业环境：激励环境

理解组织的动机和 RFID 部署的检查和分析的原因，是非常有必要的。为什么商业组织关注 RFID，以及策略和计划模型是什么？这扩展到了 RFID 实现的依据、范围和计划的范围，以及与其他识别技术的关系。换句话说，商业环境战略可能不是作为计划范围的最后结果，由于随着外部的或是内部的因素的变化，需要改变它的边界。同样的，由于整个方法三个阶段的反馈，目标可能会不断地得到提炼。如果部署过程是一堵封闭的墙，那么反馈循环将主要关注于内部的因素。在供应链部署和集成研究的案例中，很明显不同组织的相互关系在决策行为和组织的合作中将扮演主要的因素。

7.2.1.1 检查决策行为

在一个组织内决策的行为假定[2]决策的数量影响了商业中的决策过程。因此，理解为什么组织需要调查 RFID 的使用，以及激励环境是什么样的是非常重要的。接下来的图将会给出三个场景（公司 1、公司 2 和公司 3），这些场景指出了一个组织在模型上的出发点，将是如何影响他们的决策行为的。

公司 1 注意到了 RFID 技术被其他企业在供应链中应用，以及为他们的车辆在

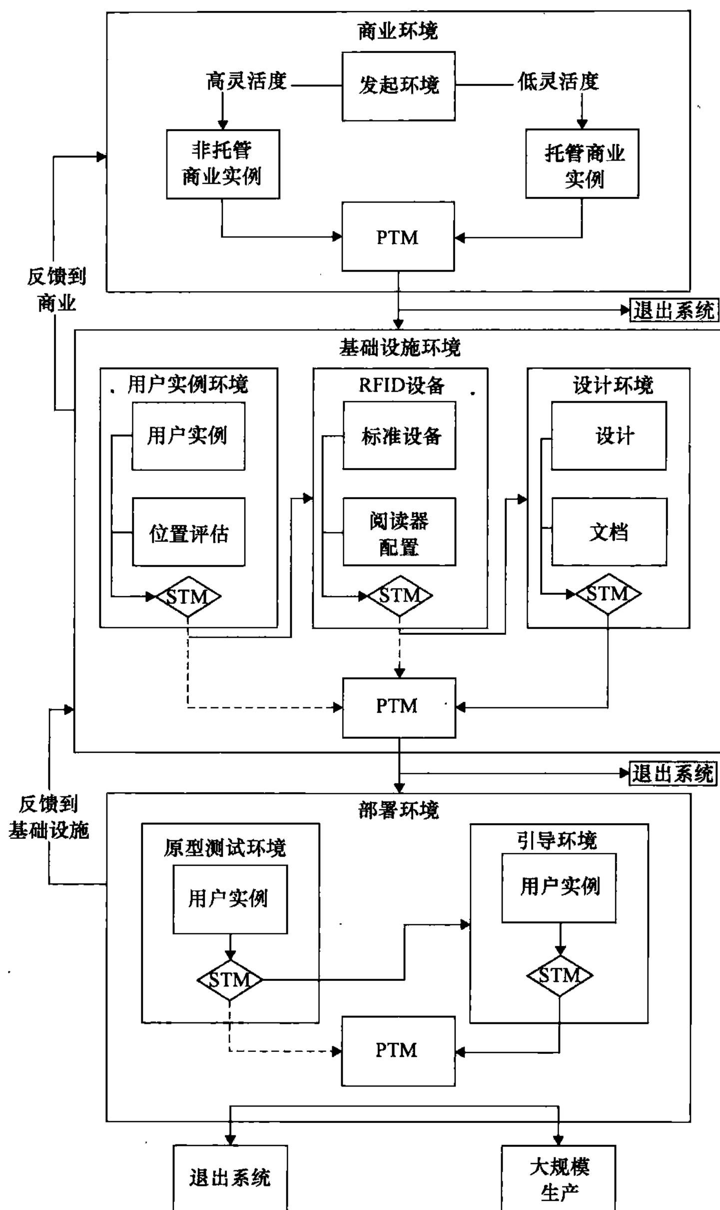


图 7-1　RFID 合理的部署模式

高速公路上行驶而购买标签[3]。他们作了一个决定，开始调查 RFID 技术的价值，以及报告一个使用 RFID 的识别系统的可行性。这个场景中，对于选择那些商业案例来探索和如何使用 RFID 系统，企业具有高度的决策权。激励环境具有激励特性，这在随后将紧跟这个阶段的非授权商业案例步骤。

公司 2 是一个主要的零售商，它提供产品给用户，它已经授权使用了 RFID 技

术，因此用户能够使用 RFID 技术来识别物品。在这个例子中，公司的决策行为被减少了，他们需要利用 RFID 技术实现特定的商业案例。虽然标签和特定的频谱必须与他们的用户系统相兼容，但这个组织仍然有能力提供他们的基础设施需求。

公司 3 与公司 2 相似，但在这个案例中，RFID 技术被授权了，并提供了实现系统的必要设备。在这个案例中，除了如何在物理上实现系统之外，组织只有极少的决策行为。如果供应商不得不允许用户意识到通过商业进行物品传递，这种自主裁量可能被进一步降低。换句话说，当物品转换状态时的能见度可能是商业案例的一个需求。这好比当产品正被制造时，制造商知道这个产品的阶段。比如，宝马（BMW）公司使用一个主动标签来识别一辆汽车在生产线上的各个阶段。标签也包含了组装汽车的一些技术规格[4]。

公司 2 和公司 3 的激励环境已经摆脱了授权方式，这导致了这个阶段的授权商业案例步骤。在这些先前的 3 个例子中，组织仍然有能力决定不用 RFID 来识别物品。如果公司决定不部署 RFID，那么在公司 2 和公司 3 的案例中，它可能有必要为了企业生存，吸引一些新用户。

表 7-1 为组织在实现 RFID 技术的商业过程中一个组织拥有的自主行为等级。这直接与激励环境相关，因为组织可能具有零到最高的自主裁量权，这取决于是否制定授权。假设在一个指定的状态中，已经决定了 RFID 频谱和标签特性。

表 7-1　决策行为

决策行为	商业案例		技术	
	授权	未授权	授权	未授权
高		×		×
中				×
低	×			
无	×		×	

7.2.1.2　工作案例研究：全国性的供应链

存在着以下的一些工作于全国性的供应链的问题。

1）激励环境是什么？

2）每个成员的自主裁量级别是什么？

3）他们工作于授权商业情况还是非授权的商业情况？

案例研究的激励环境来自于激励的基点。成员来自于非授权激励环境，并且显示了一个高等级的自主裁量行为。这些供应链的成员和其他相关利益方已经决定采用（测试）RFID 技术，从而评估该技术所带来的利益。他们对供应链中从制造商到陈列室各阶段针对物品的跟踪所带来的益处很感兴趣。

值得注意的是，商业环境模型的初始化入口具有调查特性，并且商业案例是不被授权的。很显然，这个案例研究中，成员必须选择一个商业案例，可以获取

RFID 的数据，以及分享组织间的数据。

7.2.2 商业环境：商业案例

商业案例定义为，变量支持一个方案来通过所有必要的过程和因素来实现一个特定 RFID 实现方案的商业目标。这包括了所有的资源，例如数据系统、人员以及需要的硬件。这个模型展示了激励环境下识别标签的低等级和高等级的自主授权两个路径。取决于激励环境，组织将会分析一个授权的或是非授权的商业案例。虽然需要相似的问题和分析，但在迫切决策的情况时，存在着一个差异。所有在用于辅助或者部署 RFID 系统的检查中涉及的组织，必须调查和考虑一个通常问题集。

问题的第一个阶段需要理解当前案例的商业环境。第二个阶段分析发展使用 RFID 商业案例的信息系统和数据战略需求。以下列出的问题，并没有按照重要性进行特定的排序。

1）问题的提出；来自初始化查询的额外问题。

2）给出的建议或是例子。

阶段一的问题：当前的商业环境是什么？

（1）Q1：商业是来自一个授权的还是一个非授权的激励环境？在这方面，重要的是，需要知道商业案例是否遵照以下原因进行了授权。

AQ：商业组织在商业案例或是定义整个 RFID 环境中，拥有多少自主裁量权？

这会影响决策，以及应用到商业案例分析的资源和重点的数量。商业案例是否被授权将会影响给定权重的 PTM 指标。一些 PTM 变量不会在组织授权的自主裁量行为中影响很大。例如，在一个授权环境的案例中，物品的识别性能和能见度将会比成本因素受到的影响更大）。在一个主导方授权的案例中[5]，激励的建议可能从主导方到隶属方变化。通用电气补贴了它的贸易方使用 RFID 标签的成本[6]。当描述和分析 PTM 时，这两个例子是重要的。

（2）Q2：什么最有可能改变人力资源？

在 RFID 部署于医院的案例中，工作人员抱怨额外的监视活动[7]。对于增加一个成功部署的可能性，拥有人员的支持和遵守是非常重要的。

（3）Q3：经济和商业环境影响组织的自主裁量行为到达了什么程度？

（4）Q4：商业案例与战略商业输出相联系吗？

（5）Q5：对于实现，存在时间限制吗？

与流程相关的商业案例映射、雇员的交互、机械、数据系统、供应商、内部和外部的用户以及战略商业目标相关。在商业案例中，存在着一个与数据系统的交互系统，这些系统可能是外部的或者是内部的。这些交互需要建立模型，因此通信（数据的交换）通过适当的设备，被获取和存储。在这种情况下，使用一个统一建模语言（UML）来实现对于数据系统交互的存储，通过在其他参与者和系统绑定之间显示，来形成一个重要的商业案例组件。将与数据系统交互的商业案例部分模

型化，是必要的。并不是商业案例中的全部过程都需要数据获取或者输出的功能。例如，在托盘上放置识别目标的活动，或者是把标签戴在一个病人身上，这都是不用被记录的。

接下来的阶段，两个问题检查了信息系统和数据战略。这个方法没有覆盖信息系统的集成或者是内部或外部组织系统间不同链接的表示系统的互操作。调查什么数据被获取，以及这些获取的数据如何支持商业案例，仍然是一个需要调查的关键。

阶段二的问题：什么是信息和数据战略？信息集成和数据采集需要调查和存储。

（1）Q1：这个商业案例的部署是供应链的一部分吗？

如果组织已经连入了一个供应链，对于他们来说，理解他们的商业案例和过程，以及用户和供应商使用的商业案例之间的关系，这是迫切的。存在着这样一个情况，就是在供应链中的组织已经把标签附着到了来自供应商的物品上，而这些供应商并没有使用 RFID 识别系统。

（2）Q2：什么协议用来处理不正确的 RFID 数据？

正如所看到的，RFID 很容易出现数据采集错误[8-10]。有必要减少在源处的错误可能性，以替代对这些数据采集之后再进行处理。商业组织需要意识到，RFID 容易出现错误，并且决定什么样的检测速率是一个可以接受的百分比。这可能意味着，对检测进行使用，并改变标签附着的方式。

（3）Q3：当前的信息系统是什么？

RFID 系统会产生大量的数据，这些数据需要通过一个有效的（低成本的）方式进行存储和使用。参考文献［11，12］讨论了数据流增加的数量，这就需要信息的转换，虽然参考文献［13］提出数据的转换模型可能需要增加的数据，以及这些数据将如何在数据仓库内存储。

（4）Q4：什么数据将会对供应链中所有用户都是有效的？

（5）Q5：识别的对象是什么？

商业案例已经检查了一系列的问题。这些问题需要从商业案例安静和信息系统的前景来调查。需要注意的是，先前的部分并没有产生一个全面的问题列表。将有一些组织的具体问题，这些问题在没有覆盖的商业问题的检查中，将会出现。

7.2.2.1　工作案例研究：全国性供应链

简要地说，全国性供应链（NSC）的参与者包括一个托盘提供商、包装公司、产品制造商、产品零售商、运输供应商以及硬件、软件和服务的提供商。这个团体是由来自所有参与者的代表领导的。来自先前部分接下来的问题，将在相关的工作案例研究中进行讨论。

阶段一问题：当前的商业环境是什么？

（1）Q1：商业是来自一个授权的还是一个非授权的激励环境？

NSC的成员来自非授权的环境。为了允许成员之间的交互，需要一个使用RFID过程的商业案例。一些成员也涉及了对于供应链场景不是特别重要的领域。因此，组织关于商业案例，具有一个低和高的自主裁量权。

阶段二问题：什么是信息和数据战略？

（2）Q3-4：什么数据将对供应链的所有用户有效？识别的对象是什么？

共享的数据会随着托盘和纸箱的所有权在成员之间转换而改变。识别的对象是托盘，以及纸箱级的产品。

接下来的一节，将在商业环境的背景下，检查PTM的目的和发展。它也将在NSC的背景下进行讨论。

7.2.3 商业环境：阶段的过渡动机

PTM引入组织的决策，进入到方法的下一个阶段。必须存在某种形式的度量来阐明来自分析、发展、实现或是保持组织商业案例的需求的一个输出。Myerson[1]使用一个三参数的性能指数来测量在供应链管理组织成熟模型的每个阶段的成功程度。这个方法使用了一个决策模板形成中的PTM。通过包含了对于定义PTM具有的度量和来自一个阶段到另一个阶段的转换时的输出，来扩展模型[1]。

PTM的发展将通过分析在这个方法中特定阶段或是步骤提出的问题来引导。在这个例子中，组织调查商业案例，并且思考他们为什么要部署RFID技术的原因？如果实现RFID的原因是有利的，那么组织可能会继续评估。因此组织将会探索和分析实现满足部署一个RFID系统需要的参数。接下来的图将讨论PTM相关的例子。

内部和外部的影响对于PTM的发展来说，是需要慎重考虑的。例如，来自顾客的授权可能给组织一系列的时间帧，以使其实现一个完整的RFID系统功能。在这种情况下，实现时间限制的能力可能在决定是否进入下一个阶段来说，是一个重要的因素。这种授权如果来自于一个主要的客户的话，对于组织来说，是一个基础设施的问题。在这个方法的背景下，需要问的问题是，“组织将开始分析基础设施环境了吗？”

在任何组织的经济生存中，财政考虑是一个主要的因素。取决于组织的经济位置，成本下降到一定的指标，或者满足投入目标的收益（ROI），这是很重要的。在这种情况下，财政的度量必须满足使组织能够进入到下一个阶段。PTM可能会连接到其他的PTM，例如带有一个ROI子群的财政指示器、网络代表值（NPV），以及一个具体的预算。每个因素都不得不满足PTM的状况，以使过程进入下一个阶段。

另外一个重要的考虑是当决定部署RFID技术时，法律问题的影响。由于政府规章的改变，组织可能被授权来实现这种类型的系统。在商业世界中，保持竞争的优势可能是调查这个技术的另一个原因。其他可能的PTM也可能会包括人类的或

是技术上的限制，业务限制，以及在供应链各成员之间的信息共享，或者是与存在的组织过程和信息系统集成。

另外需要考虑的一点是，一个授权的或是非授权的商业案例的结果是如何影响 PTM 的关系和发展的。如果企业是来自于一个授权的商业案例，那么它们可能会显示出对于 PTM 结果的低耦合。如果授权是来自于与销售量有关，且较有影响力的用户，那么这对于供应链管理中的组织来说是恰当的。如果组织希望继续它们当前的商业环境，那么组织可能需要实现使用 RFID 的商业案例。因此，PTM 步骤的结果可能并没有与财政利益密切联系。例如，NPV 或者是 ROI 可能并没有像组织实现这个授权的需要一样有影响。另外，非授权的商业案例可能与 PTM 的输出存在着一个较高的耦合。例如，ROI 将与持续 RFID 部署存在一个较强的相关性。

在授权的商业案例中，无条件的行为将会减少，这可能会影响 PTM 参数的设置。在这样的场景中，有必要把商业环境的目标和 PTM 连接在一起。如果目标没有符合，那么商业案例和 RFID 部署的一个进一步的相关，将是有必要的。一旦 PTM 发展，分类并且给予了一个想得到的结果，例如一个基于货币或者时间的目标，那么它随后将会进入到一个 PTM 的模板中。见表 7-2，为方法的每个阶段或步骤的 PTM 速率系统。

表 7-2　PTM 模板

PTM			
加权	可行性	结果	行为
紧急的	可行的	成功	连续的
被推荐的	挑战性的	再评估	反馈
最佳的	不可行的	失败	退出

从表 7-2 可以看出，一个权重被分配到了一个 PTM 或是步骤转换动机（STM）中。PTM 与一个阶段相关，而 STM 则与一个阶段中的其中一个内部步骤相关。为了简化阅读这些关系，PTM 将在本部分使用。每个具有的 PTM 将会有一个可能性的规模，这与组织完成 PTM 的能力相关。一旦 PTM 的分析进行，一个结果比率被分配到特定的 PTM 内。这会在一个阶段或是一个步骤完成之后发生。根据这个结果，组织将决定一个动作。这个动作可能会继续评估的过程，产生反馈给一个先前的阶段，或者是退出这个方法。在最后的情况时，这会与检查 RFID 部署的终止相关。

在表 7-3 中，财政考虑给出了一个必要的权重。换句话说，财政度量必须符合，否则组织将不能继续 RFID 的部署。在这种情况下，商业案例不是授权的，财政结果对于决策过程来说，是较为重要的。所有的 PTM 在方法中完成一个反复循环。表 7-3 描述了，如果第一个循环结果是失败的，那么反馈将给到先前的阶段或是存在的阶段中。随后，如果下一个反复结果是失败的，那么将产生“退出”

动作，组织将会终止部署。如果动机权重是必要的，那么这当然会是一种情况。另外一个例子也可以在表7-3中看到，其中，组织已经决定了它想由本地的设备提供资源。这个权重是建议的，并且是一个“继续”的动作，即使结果是失败的。因为组织不能够在本地获得设备，这个PTM的结果将不会停止基础设施阶段的过程。

表7-3 财政的和RFID供应商PTM

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	财政的	紧急的	可行的	失败	反馈
2	财政的	紧急的	挑战性的	失败	退出
1	局部RFID供应商	被推荐的	挑战性的	失败	反馈
2	局部RFID供应商	被推荐的	挑战性的	失败	继续

方法的这个阶段，将会是组织接下去的步骤数量的一个决定因素。换句话说，如果PTM没有符合组织的要求，那么RFID部署的调查可能会退出。

7.2.3.1 工作案例研究：全国性供应链

由于本书篇幅的限制，对工作案例研究中的PTM或是STM的讨论，将只给出一个例子。在这个例子中，组织调查了决定，来继续基础设施环境阶段。见表7-4，商业环境考虑的法律顾虑，是与PTM有关的。由于关键是要符合所有的标准，所有它被给定了一个必需的权重。这个可行性确定在部署的这个阶段是可行的。结果是成功的，PTM的动作将持续进入到下一个阶段。

表7-4 法律上考量PTM

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	合法的	紧急的	可行的	成功	继续

商业环境决定了一个商业案例的目标和范围。下一个步骤将是确定PTM，并分析他们的结果，决定一个动作。如果结果是需要持续的，那么他们将会继续进入到下一个阶段。这个阶段检查了基础设施环境，决定在一个业务设定背景下商业案例的可行性。

7.3 第二阶段：基础设施环境：制造商到零售商

RFID解释和部署方法的这个阶段调查了基础设施环境，以及主要根据RFID物理特性分析的一个关键成分。这一理解将会映射到商业案例，以及随后的使用案例中，可以增强或是开发使用RFID的过程。射频信号易受电磁场的干扰而亏损。

如图 7-1 所示，为部署的这一个阶段，组织调查哪里和怎样完成这个使用案例。现场评估建立了一个可能干扰的画面，其中使用案例被执行，并把信息输入到了 RFID 设备步骤中。STM 是在这个阶段使用案例环境、RFID 设备和设计片段中的每个步骤的一个评估检查点。一旦使用案例环境的 STM 度量被计算，并且与这个激励阶段的下一个步骤组织的需求一致。第二个步骤是 RFID 设备，并涉及 RFID 阅读器和标签的设置，以便于在使用案例环境其他活动的最小化干扰的情况下，优化对于一个对象的识别。设计步骤进一步发展了来自 RFID 设备环境的输入。这个步骤是为了环境中的每个使用案例和全部的活动而完成，这需要与对象进行交互。一旦这个阶段的 PTM 检查完成，组织可能会进入到部署阶段。

7.3.1 使用案例环境

使用案例环境由一个作为商业案例部件的使用案例，以及每个使用案例的现场评估组成。使用案例主要是与一个组织设置中的一个对象的识别相关。这将会在接下来的部分进一步的解释。

7.3.1.1 使用案例

来自商业环境的每个商业案例的细节是建立这个阶段的基础。在先前阶段完成的分析和存储，将会设置这个步骤的方向。使用案例环境与在一个特定的商业设置中一个附着标签的对象和一个阅读器的交互所采集的数据相关。这就是为什么有必要具有一个清晰的商业案例的书面文档知识，以及它的目标和范围。这个知识提供了额外的视角，这有关于使用案例将如何使用 RFID 技术。这主要的目标是记录数据通信将会如何以及在哪里发生，以便于现场评估将具有一个初步的位置蓝图。使用案例环境的进一步评估能够从仔细的观察，与雇员的会面，文档记录和系统识别器的先前经验中得到。

这个步骤需要定义一些问题，以及分析的结果来帮助使用案例的理解。

需要回答的典型问题包括下面一些例子。

(1) Q1：使用案例是属于什么商业案例？

AQ：使用案例的主要目标和范围是什么？

了解商业案例，以及使用案例是怎样映射到商业案例中，是需要慎重考虑的。这个使用案例可能是许多使用案例中的一个子集，这些使用案例需要来满足一个特定商业案例的功能。了解使用案例之间的联系将会对每个数据交互的范围和目标的理解有较大的帮助。这就是为什么了解一个特殊的使用案例的情境是非常重要的。

一个例子就是一个托盘上进行按顺序的包装。

首先，使用事件 1，空的托盘会被检索；使用事件 2，产品用一个纸箱包装；纸箱会被放入一个托盘；使用事件 3，托盘就被包裹，然后这 3 个可能会循环，直到排序完成为止。

使用事件 4，如果运输准备好时，托盘将会放入一个装载区。使用事件 5，或

者托盘会放到竞争顺序区域。

这就是排序过程商业案例的范围。每个使用案例会在一个概念模型中被存储和测试，也和在基础设施环境中阅读器和附着标签的物品之间的交互有关。一些有用的模型模板是与活动图表相邻的使用案例描述和图表。图 7-2 是一个使用案例的图标，显示了与系统交互的一些角色。这包括了一个包裹站，一个雇员，以及带有一个 RFID 阅读器的叉车。当叉车提取一个空的托盘时，由于托盘数量的减少，系统会注意到，因为雇员完成包装附着标签以及记录标签 ID 号的纸箱。托盘随后会用塑料包裹，并准备运输。

（2）Q2：使用案例的数据需求是什么？

AQ：什么数据将会被发送，或是写入标签？

AQ：信息系统需要什么数据参数？

正如图 7-2 所示的，提取空托盘需要的数据是托盘的标签 ID 号。在这种情况中，不需要数据写入到标签。在托盘包裹站，托盘的标签会被记录，每个纸箱的标签 ID 也会被记录，并且带有每个使用案例的一个时间戳。根据组织的使用案例场景，每次读的位置也会被记录。

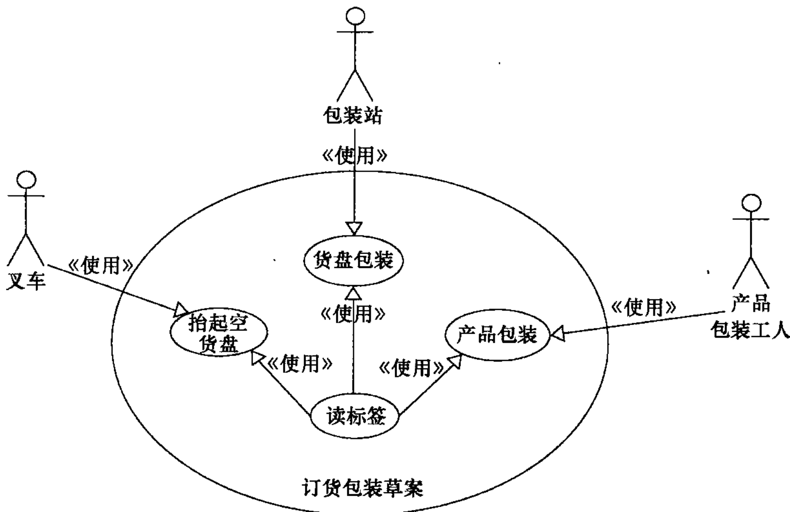


图 7-2　产品订货的使用案例

（3）Q3：使用案例和识别区域的位置在哪里？

AQ：RFID 的频谱是什么？适合使用案例的 RFID 的频谱和标签类型是什么？

AQ：环境电磁噪声是什么？

AQ：如果物体没有被识别，那位于什么系统中？

AQ：这个使用案例是供应链中的一个链接吗？

这个步骤结果的存储需要指定一个格式，这个格式是被组织和部署小组所理解

的。统一模型语言（UML）模型被称为一种结合使用案例描述的动态图表，它对于将需要的RFID数据传输的存储过程来说，是一个有用的工具。这些模型也将帮助在每次数据交换前后发生什么活动。这些知识能够帮助优化每个RFID识别区域的位置，突出每个RFID系统需要的数据和识别物体的状态由于这个交互，是怎样改变的。信息系统也需要信息，以及数据将在哪里存储和怎样存储这些数据，来自这个过程的数据需要什么系统，以及这些数据会存在哪里？

一个使用案例的基本需求是在一个RFID系统中对物体的识别。因此，从这个阶段中，一个组织可能能够建立有一个基本RFID过程的模板。一些在供应链中识别的基本例子包括：在门口识别物体；在包装站识别物体；在传输阶段识别物体；在查询位置处识别物体；以及在组装阶段识别物体。

在参考文献［14］的调查中，参与人员在防御部门（DOD）供应链中识别作为RFID技术实现的主要范围的传输。物品级别跟踪的使用仍然是一个扩展和实现的区域。他们也观察到，识别数据的被动标签是主要的RFID标签。一个通用模板的发展，将帮助创建使用案例分析和文档的一个标准过程。例如，组织可能在一个位置调查RFID，这些模板将会帮助在可测量的范围具有相似需求的另一个地理设置的可扩展性。表7-5所示为这个模板，包括了来自使用案例环境，RFID设备以及设计环境中作为一个基础使用的信息。利用全部的模板，改变可能需要满足一个具体组织部署的需求。模板基于一个使用案例描述，虽然它增加了与RFID阅读器，天线和标签等相关的类型和放置位置的信息。当更多的信息在基础设施环境步骤结束期间有效时，模板是可以被更新的。

工作案例研究：全国性的供应链

接下来将讨论来自先前章节的关于工作案例研究的问题。

（1）Q1：使用案例属于什么商业案例？

AQ：使用案例调查了什么？

这个问题完全取决于每个成员。一个成员调查商业进出的托盘数量，其他成员则检查一个顺序上的包装和运输。一般地，使用案例是在组织的背景下与一个对象的识别相关的。使用案例在供应链的每个实体之间变化。工作环境从托盘供应商到制造商、运输公司、产品入仓库以及产品到商店变化。

表7-5　使用案例模板

使用案例	包装托盘			
组织	公司×			
位置	包装站3建筑B（见计划B）			
操作环境	室内、低温度、包装站靠近建筑物需要的进入点			
地址允许	来自场地评价的结果			
对象类型	托盘	纸箱	产品	物品

（续）

对象标签	EPC 被动 RO	EPC 被动 RO	无标签	
系统数据	时间，位置	时间，位置	无标签	
标签放置				
成功读取（每次）				
操作范围	标签到阅读器			
频率	UHF 918～926MHz			
阅读器	模型 X（配有 IP）手持/4 端口			
阅读器放置	待决定			
天线	A1	A2	A3	A4
天线放置	待决定			
预先准备	建立好顺序，并且托盘已经被打包，现在开始包装托盘			
后置	托盘被包装好并识别，然后移到装载区域			
例外	托盘中的物品丢失			
注意	标签有一个预设的 EPC ID 号，对象的状态是等待装船			
附加要求	阅读器入口需要铃声			

（2）Q2：使用案例需要什么数据？

所有供应链中的参与者想要看到物品从制造商到陈列室的移动。这意味着获取和共享例如物品的时间和位置这样的数据。这个知识需要提供正确的数据给供应链上下游的成员。例如，托盘提供商想要知道托盘在供应链中的移动。因此，包括读托盘识别的托盘的接受者在读取中包括托盘识别，并且在 web 共享入口处可见都是非常重要的。

一旦使用案例被存储，现场评估会被建议分析组织的使用案例环境。这将在接下来的部分进一步地讨论。

7.3.1.2　现场评估

现场评估将涉及在一个给定使用案例的商业循环上的一个完全的法拉第循环分析（FFCA）。理解可能物体或许会在标签和阅读器之间通信造成干扰是很重要的。这个干扰可能由下列一些原因引起：例如报警系统的 RF 通信设备；无绳电话；以及在建筑结构中使用的例如金属框架这样的材料。

实体和部署实体的一个重要问题是："是否有活动在一个不可推测或是不经常的时间循环内执行？"例如，使用 RF 设备的通信员不经常地采集具体的物体。这些问题证实，现场评估必须在一个完整的商业循环中执行。这需要在整个商业过程中为周围的 AEN 映射环境。在前面的案例中，评估小组必须意识到在整个商业循环中例外的环境，以便于能够获取可能的阅读器到标签干扰的全部例子。

执行一个现场评估主要的原因是为了调查部署的网络和附着标签物体的环境。在这个过程期间，有必要绘制商业过程，因此，需要传输或是 RFID 数据可能传输的位置将会被存储。这些需要决定了每个位置和部署区域蓝图创建的识别区域。客户和系统集成商需要读取区域放置的具体信息，因此，蓝图对于成功的安装是重要的。这个返回的信息允许客户知道商业过程的可能改变，以最大化标签附着物体的识别。

下面的图将给出一个执行现场评估过程的概述。将讨论的主题是：一个 FFCA 需要的设备；FFCA 是如何被组织的；标准配置参数是什么；以及这些调查的结果是什么。

FFCA 必要的设备包括如下：频谱分析仪；带有接地的 1/4 波长偶极 992MHz 天线，这个天线可以 360°获取连接到频谱仪的电磁波；连接到频谱分析的记录设备。

监测区域将设置专门设备，读取的数据将在需要的使用案例时间循环上被记录。频谱分析仪设置为使用国家 UHF 频率的中心位置频率。天线被放置在 IZ 的中心，这将取决于每个现场场景。记录设备被设置为登记来自频谱分析仪的在一个设置的间隔或是在使用案例时间循环期间具体商业操作的数据。从这个测量过程，有必要对许多读数据的区域重复这个测试，以便于跟踪阅读器和标签通信的潜在干扰。这取决于在 IZ 分析和 IZ 场景期间采集的噪声数据。具有全部的可能造成 AEN 通过识别区的机制是非常重要的，因此数据可能会在它的工作噪声上进行采集。如果只有一个传送带组装线被测量，那么这是不可能的。因此，可能许多场景需要来调查，以决定 AEN 的产品。

当这个操作进行时，所有的噪声源需要在部署的蓝图上被识别，以及把细节记录到使用案例模板中。这个蓝图是这个活动的主要输出，并带有 IZ 位置放置的建议。从现场评估，将可以清晰地看到，RFID 的性能将是什么，以及在设备选择，被指和位置安装上的引导。必须指出的是每个现场都需要分析，并且所有现场都不同。

工作案例研究：全国性的供应链

现场评估的结果是什么?

可以观察到每个组织都有具体的现场问题。不可能为将来自一个特定的现场评估结果应用到另外一个现场评估而回执一个通用的结论。一个成员存在来自射频系统的干扰问题。这可以通过为托盘顺序识别创建一个时间帧来予以克服。

现场有运行在 950MHz 频率的已存在系统，并且需要注意到可能会出现新的 RFID 系统的干扰。分析完成，结果报告认为应当是没有问题的。

7.3.1.3　使用案例环境：步骤转换动机

使用案例环境的 STM 设置的执行是为了度量这个步骤的有效性。建议动机的接下来的清单可能是一个组织的开始点。有必要了解这个阶段的下一个步骤是否将

开始。

（1）STM1：AEN的等级是什么？

这将会有一个必要的权重。这个动机对于RFID系统的操作是有必要的。如果AEN过高，在一个IZ内的阅读器和标签之间的通信可能会不工作。

（2）STM2：使用案例重复引擎

初始化时，使用案例重复引擎将有一个可选的权重，但是由于在AEN高等级的情况下，如果IZ的位置需要移动，那么它可能变得是必要的。可能需要调整现场评估的结构的适应性，以便于优化阅读器和标签之间的通信。一旦这个阶段的RFID设备步骤完成，进一步的信息将是有效的。

如果商业环境PTM包括了一个必要的时间限制或者是预算限制，那么这些将需要在这个步骤中考虑。如果来自这个使用案例环境的结果对实体是有利的，那么基础设施环境的下一步，即RFID设备环境会被初始化。如果全部必要的STM有一个持续的动作，那么组织将开始RFID设备步骤。如果是“退出”动作，那么这个步骤将携带结果进入这个阶段的PTM过程。

工作案例研究：全国性的供应链

由于空间限制，这个阶段的一个STM或是PTM将会研究与方法相关的NSC例子。

这个STM（见表7-6）与引起已存在系统的频率干扰的RFID部署的推断相关。在这个现场的评估之后，经决定将不会有令人担心的结果。

表7-6　AEN层STM

STM					
迭代	激发因素	加权	可行性	结果	行为
1	AEN层	紧急的	挑战性的	重评估	反馈
2	AEN层	紧急的	挑战性的	成功	继续

7.3.2　RFID设备环境

RFID设备环境调查实现一个RFID部署需要的硬件和软件。进行阅读器和标签通信的测试和配置，用来得到一个优化的识别速率。标准设备和阅读器配置将在接下来的几部分进行讨论。

7.3.2.1　标准设备

质量保证意味着RFID设备需要在部署之前进行检查。这个设备检查阅读器和天线运行的功率和频谱覆盖范围在国家指南以内，并且是可使用的。标签的测试能够在一个被控制的测试实验室内完成。这对于决定最小化的功率是有必要的，需要启动标签，并获得一个平均的标签度量。购买优化的设备，以便于适应需要的使用案例，这也是很重要的。在一个供应链中，一个组织的RFID系统必须能够与用户

或是供应系统通信。正如之前提及的，射频频谱常常是应用相关的。在设备购买之前，互操作问题相关的信息需要避免，这是很有必要的。RFID 之间不同的标准和不同的国家频谱参数可能会造成在 RFID 部署时的问题。需要意识到应该形成一个有效的系统，以便于兼容组织将来的扩展，或者可能与其他组织的交互。对于网络阅读器的低级阅读器协议（LLRP）的发展标准，将会在 RFID 技术中解决一些互操作问题。使用案例模板帮助选择 RFID 设备。这种特定的模板确保了每个读范围中，所有必要的组件都可用。这包括了阅读器、天线以及不间断电源（UPS）等。

标准的使用将帮助组织产生一个通用的使用案例，因此它可以被部署到其他的组织设置中。这个标准需要：测试过程；文档；具体的设备；频谱参数；采购设备过程，以及数据传输和格式。

这些协议将会指导 RFID 部署所有相关的成员，无论他们是一个供应链或者闭合环路系统的一部分。

工作案例研究：全国性的供应链

澳大利亚的 UHF 标准是试验中使用的标准。它的频谱为 918 ~ 928MHz，天线的功率为 1W。使用的标签是 class 1 类型的被动标签。阅读器为固定式的或是手持式的。固定式的阅读器使用一个 IP 地址连接到网络，并且能够使用多达 4 个天线进行通信。合作团体通过工作案例研究的设备供应商成员了提供阅读器和标签。这确保了所有的成员可以从同一个供应商中获得相似的设备。

7.3.2.2 阅读器配置

阅读器以及它们的天线的配置，对于标签的识别是非常重要的。一个 RF 路径衰减轮廓图的调查的执行是为了优化阅读器到标签通信的形成。阅读器和它标签的信号将被调查，来识别标签是否将被识别。了解每个 IZ 的边界将弄清天线覆盖和标签识别所需要的距离。为了克服交叉读取，天线功率的减少可能是有必要的。或者，如果 IZ 一边的干扰影响了读数据，那么另一边的天线可能不得不增加它们的功率。

在配置期间，注意天线和阅读器的位置是重要的，这个位置会由于现场限制和干扰而改变。天线实际的位置将包括它们到标签和 IZ 的角度。阅读器的配置包括以下一些内容：IP 地址；天线端；读循环；标签采样速率；以及到控制器和中间件的连接。本章没有考虑阅读器和组织系统的交互和集成。

工作案例研究：全国性的供应链

在读湿的或者是刚上油漆的托盘时，存在着一些问题。这个识别问题可以通过当标签贴附到托盘时，在标签后使用泡沫材料和铝箔来予以减少。

7.3.2.3 RFID 设备：步骤转换动机

执行对于 RFID 设备的 STM 发展，以分析这个步骤的结果。建议动机的下列清单是组织可能使用 STM 的一个例子。

1）STM1：成功的物体识别的百分比

应用到这个STM的权重是必要的，虽然也可能需要一个二次识别系统。

2）STM2：优化标签位置和方向

这是与STM1相关的，对于成功读数据将会有重要的影响。另外一个联系的动机是标签到阅读器通信的配置。

工作案例研究：全国性的供应链

在阅读器配置期间，会发现物理端口数量并不与阅读器的软件端口数量一致。阅读器被设置读0到3端口的数据，但是它却试图读1到4端口。这是一个标准问题，并且决定继续部署。这个问题强调了在设备部署之前测试这些设备的需要（见表7-7）。

表7-7　阅读器标准STM

STM					
迭代	激发因素	加权	可行性	结果	行为
1	阅读器标准	被推荐的	可行的	重评估	反馈
2	阅读器标准	被推荐的	挑战性的	重评估	继续

7.3.3　设计环境

这个在基础设施环境的步骤发展了满足一个使用案例的目标和范围需要的设计。来自这个阶段的过程步骤的输入作为这个步骤的基础使用。输出是RFID系统部署有关的文档的采集。

7.3.3.1　设计

设计是用来发展基础设施的蓝图，以使RFID技术在一个使用案例中使用。设计步骤是一个RFID部署的实现方案和操作文档。设计一个RFID部署需要阅读器、天线和数据需求的实际位置的信息。这包括了为放置电缆、不间断电源、设备位置以及面积网络连接点创建一个方案。例如传动带状况、托盘湾、在现场评估中来分析的阅读器和天线的码头门这样的现场场景预先安装的基线配置需要额外的设置来创建。设计步骤也产生了信息系统和服务器需求相关的文档，例如包含阅读器，天线和物品上标签的防治需要什么样的现场基础设施。

来自使用案例环境的文档，RFID设备以及商业环境在方法的这个阶段实现。一个必要的输入是使用案例模板，它可能对一个位置包装站、产品架或者是防盗（在零售店的入口处）的物品的识别产生文档。模板参数可能需要对每个实现进行调整操作，作为一个特定的现场，可能推断，这不会在另一个现场发生。

当设计一个商业过程的安装过程和设备图时，重要的是最小化对工作流的分解。例如，阅读器和天线的位置，由于RFID基础设施的放置，可能需要为车辆而改变方向。一般地，阅读器站有必要安装保护外壳，因为阅读器站可能被某些机械撞击。一个设计完成，文档步骤也会被执行。

工作案例研究：全国性的供应链

需要注意的是完整的理解使用案例，以及它如何结合 RFID 设计计划来工作。用户接口的发展以及新的过程必须考虑到个人。

7.3.3.2 文档

文档包含了一个到多个使用案例部署的说明和协议。它也集中存放了在这个阶段产生的所有文档。它有效地把文档划分为业务，设计，安装分割，因为并不是所有的文档都将会被参与者在系统的实现中使用。文档的例子包括：员工的培训材料；一个完整的材料清单；部署的测试过程；风险评估；部署相关的人力资源考虑，以及用户接受测试协议。

文档的质量保证和反馈到使用案例的设计参数映射是这个阶段的一个重要过程。这个映射将会帮助找到可能错过的或是没有正确理解的信息。当设计一个供应链部署时，这尤其重要。

工作案例研究：全国性供应链

拥有正确无误的使用案例文档以及这些文档与组织的交互非常重要。由于这是一个供应链，有必要具有对全部参与者需要的数据和范围的一个清晰的理解。

7.3.3.3 设计：步骤转换动机

设计步骤动机是否完成这个阶段到基础设施环境的 PTM 步骤的过程。可能的动机包括：

1）STM1：法则的遵守。

这是必要的，因为你不想不服从，例如，在电缆或者是部署协议发生时。

2）STM2：正确无误的文档。

这是一个必要的 STM，必须与组织商业案例需求进行映射。重要的是拥有一个清晰的使用案例文档，以及它的发展，以及与现存的或是新的系统的集成。这取决于场景中一个建议权重的应用。

工作案例研究：全国性的供应链

设计步骤的一个 STM 的例子是准确文档的产生（见表 7-8）。清楚地看到，在不同的企业之间共享了足够的信息。

表 7-8 精确的文件 STM

STM					
迭代	激发因素	加权	可行性	结果	行为
1	精确的文件	被推荐的	可行的	成功	继续

7.3.4 基础设施环境：阶段转换动机

这个阶段的 PTM 是部署环境的关口，只有来自动机的动作被继续。如果动作被反馈，那么这可能间隔商业环境（这可能是商业环境内部的）。如果这是一个第

二次的交互，并且是个“退出”的动作，随后组织可能终止调查。可能的PTM例子包括以下内容。

1）PTM1：系统对规定的服从

这与政府、工作地点安全规定和RFID设备相关。给定一个必要的权重，系统必须服从，否则可能会发生经济处罚。

2）PTM2：使用案例到商业案例的映射

这个阶段分析了与一个与使用案例相关的商业案例的范围和目标。

7.3.4.1　工作案例研究：全国性的供应链

一个可能应用到NSC的动机是IZ的位置，见表7-9。用它来显示物品状态的改变，这是很重要的。换句话说，在一定识别区域内一个物品的识别产生了物品所有权的改变。

现在，既然基础设施环境已经完成，组织可以进入到这个方法的部署环境阶段。

表7-9　询问区

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	询问区的位置	紧急的	挑战性的	重评估	反馈
2	询问区的位置	紧急的	挑战性的	成功	继续

7.4　第三阶段：部署环境：工厂到陈列室

部署关键检查运行一个使用RFID技术的使用案例过程的结果。正如图7-1所示，这个阶段具有两个不同的过程：原型测试以及试验环境。原型测试首先完成，以及如果（是否）试验被委托成功。

7.4.1　原型测试环境

原型测试环境检查是一个商业案例部分的使用案例。如果STM的结果具有一个持续的动作，那么组织将开始一个试验研究。否则，方法过程会进入PTM步骤。

7.4.1.1　使用案例

正如商业案例部分提及的，以及在使用案例的产生顺序图（见图7-2）中显示的，一个企业案例可能有一个到多个的使用案例。原型测试评判需要完成一个企业案例范围和目标的使用案例的集成。有必要测试这些使用案例，以便于他们能完成一个企业案例的问题。单个使用案例IZ在基础设施环境阶段期间已经被测试和设

计。在那个阶段，开发和安装过程经过细致的发展，被使用以实现这个使用案例。这个阶段的一个目标，是理解在一个业务（运行）状态期间，一个使用案例是如何运行的。

阅读器、天线和标签的配置已经形成了文档。需要强调使用案例的范围运行在它需要的等级。例如，在使用案例图（见图7-2）中，一个托盘的内容会在一个包装站识别。这个场景使用了 RFID 技术，并且结果会与使用案例操作度量进行比较。这些度量能包括：识别速率；获得的正确数据元素；正常运行时间；运行成本；组织信息集成。一旦这个使用案例生效，那么下一个使用会被测试。这将会持续进行，直到每个使用案例的性能测量被测试，以及与企业案例相关成功为止。

在供应链案例中，每个组织将会测试一个范围的同意的企业案例。它们将测试一个过程，这个过程允许指定的物品被识别，而在供应链中的连接之间移动。一个例子是对托盘和纸箱级别物品的跟踪。只有一些成员将需要识别全部的物品，例如在供应链交互作用期间的纸箱和托盘。因此，为了使这个物品产生可视化，每个组织将测试需要的使用案例，来执行这个活动。一旦需要的使用案例在每个组织中运行，那么试验能够开始。

这个模块化测试原型的方法，将允许组织去观察它是否可能实现一个选定的企业案例的试验。在组织资源没有重大破坏和延伸的情况下，组织可以得到每个使用案例可能输出结果的一个图。这包括了改变管理方案，即变动逐渐被继承到组织的运行中。结果的评价和学习的经验能够记录到问题的下一个片段中。在试验开始之前，这产生了一个 RFID 基于可能好处的细节评估。一旦选择的企业案例被验证了，现在它将会准备在一个试验部署中使用。

工作案例研究：全国性的供应链

虽然 NSC 没有一个设计好的原型测试环境步骤，重要的是在部署一个试验的 RFID 系统来克服没有文档的问题之前持续测试。例如，当新的 RFID 频率被部署时，明显地，它将会与现存的铲车系统产生干扰。这是先前测试过的，结果显示这不存在问题。从这里得到的经验是，在部署之前需要持续进行测试。

7.4.1.2　原型测试环境：步骤转换动机

原型测试环境 STM 的设置将进行这个步骤有效性的度量。建议动机包括以下部分。

1）STM1：使用案例集成。

这将会有一个必要的权重。这个动机对于 RFID 系统的运行是重要的。

2）STM2：人员培训。

这个 STM 有一个必要的权重，并且将提高试验系统的接受性和成功。

工作案例研究：全国性的供应链

按照方法，把一个 STM 应用到 NSC（见表7-10），在试验前对使用案例持续的测试，将克服任何意料之外的问题，特别是铲车系统的干扰。

表7-10 用户实例测试STM

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	用户实例测试	紧急的	挑战性的	重评估	反馈
2	用户实例测试	紧急的	挑战性的	成功	继续

7.4.2 试验环境

一旦选择的场景被各方批准，那么一个计划良好的、可升级的、域内的试验研究将会开始。在一个完整的集成被完成或考虑之前，试验将允许RFID系统能力的进一步调查和确定。

7.4.2.1 使用案例

使用案例的识别将来自原型测试环境的输出。这可能是大多数决定的有前途的完成范围企业案例的一个授权。试验的目的是运行一个更长的时间帧，来确定先前步骤的结果。

应急必须到位，以便于克服可能产生的困难，特别是系统停机，以及错过或是不正确的数据获取。当实验过程进行时，标签放置和工作寿命将变得明显。由于物品的处理，标签可能需要移动或者重新放置。例如，标签附着到某一特定的物体，可能会随着时间的推移而脱落，因此应当标明一个可以代替的标签类型的来源。

数据流的整体增长，以及存储和转换这个数据流为有用信息的需要将变得更加显而易见。特别是在供应链案例中。由于在一个IZ中交叉阅读，一个特定的参与者可能会上传不正确的数据。这就有必要对数据源进行识别，然后对识别区重新配置。

度量测试与用户接受一个新的软件产品的测试相似。用户组织签署了一个协议，即如果条件满足，那么部署的那一部分是成功的。这可能是一个室内安装或者是由一个第三方组织的。一个例子是精确速率，在一个给定区域成功读的次数，以及测试将如何运行所需要决定的参数。例如，客户希望读托盘上的30个纸箱，这些纸箱随托盘一起通过一个闸门。需要的数据准确率是99.5%，测试将会进行10次。拥有一个测量参数是很重要的，这会在一个特定的场景中，给双方签署一个里程碑。一旦这已经发生了，那么一个可能的试验研究可能会开始。

工作案例研究：全国性的供应链

NSC试验是在供应链中跟踪期望目标能力的一个成功描述。参与者能够比使用传统的条形码更有效地提高托盘和纸箱的识别和可视度。共享信息实时有效，这增强了组织的决策。一些重要的经验包括：完全的范围和备份的使用案例的存档；试验的内部和外部的组织支持；有经验的RFID合作方的使用；使用案例持续的测试和改进；缓慢的扩展RFID的实现；以及在RFID安装之前，基线度量的发展。一些难点包括：缺乏现存的标准，造成了设备互操作的限制；来自现存系统的干扰；

以及错误数据的数量。

7.4.2.2　试验环境：步骤转换动机

试验环境的 STM 的设置将进行度量这个步骤的有效性。建议动机的下列清单可能是一个组织的开始点。

1）STM1：里程碑的决定。

例如，签订的协议将制约那个使用案例的度量分配。清晰的目标和性能测试是组织之间的协议构成所必需的。

2）STM2：改变管理性能。

如果来自这个 STM 的结果不是明确的，那么成功试验的可能性将会被破坏。

工作案例研究：全国性的供应链

NSC 的试验环境的一个 STM 的例子对于人员成功的培训是重要的（见表 7-11）。没有在工作实践中改变的理解，案例研究将不能成功地执行试验。

表 7-11　雇员训练 STM

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	雇员训练	紧急的	挑战性的	重评估	反馈

7.4.3　部署环境：阶段转换动机

进行大量生产或是退出方法的转换动机是什么？在企业环境中发展的 PTM 可能仍然在现在和先前的阶段存在一个承载（一定影响）。这对时间限制，经济考虑，规定服从和运行考虑是有效的。

1）PTM1：企业案例范围和目标的完成。

如果这个 PTM 成功了，那么这个企业范围将变成一个生产单元。

2）PTM2：供应链协议。

如果所有各方同意这个特定的试验是成功的，那么它可能会成为一个生产单元。这将给定一个可选的权重，取决于它是否是一个非授权的方法。

7.4.3.1　工作案例研究：全国性的供应链

部署环境的 PTM 的一个可能例子是在所有各方之间的协议，或是它是否变成一个授权的企业案例（见表 7-12）。NSC 是一个试验研究，具有使用一个被定义的寿命来描述一个 RFID 供应链的能力。

表 7-12　命令商业实例 STM

PTM					
迭代	激发因素	加权	可行性	结果	行为
1	命令	紧急的	挑战性的	重评估	反馈
2	命令	紧急的	挑战性的	成功	继续

7.5 结论

RFID技术的分析、计划、协调、测试、逻辑依据的文档以及部署，是RFID成功集成到组织中的至关重要的步骤。本章概述了使用一种全国性供应链案例研究来作为一个RFID部署的方法切入点。激励环境将会影响一个实体的无条件行为，以及它如何选择部署的方法。分析和圈定商业环境是非常重要的，因此需要定义一个清晰的和性能量化机制。基础设施环境检查了必须被理解的RFID的物理和技术需求，以及通过RFID过程来加强或者发展的使用案例。一个可能成功的关键是对所有参与人员的协商和训练，这就需要初始化改变一个管理过程。对实现环境的完全理解和试验研究能够使商业估算出RFID部署所形成的影响。

参考文献

1. Myerson, J. 2007. *RFID Business Processes RFID in the Supply Chain: A Guide to Selection and Implementation.* Boca Raton, FL: Auerbach Publications.
2. Williamson, O. E. 1964. *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm.* Englewood Cliffs, NJ: Prentice-Hall.
3. Landt, J. 2001. Shrouds of time—The history of RFID. Tech. rep., AIM Inc.
4. Poirier, C. and McCollum, D. 2006. *Chapter 9—Building the RFID Business Case and Roadmap for Execution RFID Strategic Implementation and ROI: A Practical Roadmap to Success.* J. Fort Lauderdale: Ross Publishing.
5. Whitaker, J., Mithas, S., and Krishnan, M. 2007. A field study of RFID deployment and return expectations, *Production and Operations Management,* 16(5): 599–612.
6. Lucas, H. 2005. *Information Technology: Strategic Decision Making for Managers.* Hoboken, NJ: Wiley.
7. Fisher, J. and Monahan, T. 2008. Tracking the social dimensions of RFID systems in hospitals. *International Journal of Medical Informatics* 77(3): 176–183.
8. Wang, F. and Liu, P. 2005. Temporal management of RFID data, in *VLDB 05: Proceedings of the 31st International Conference on Very Large Data Bases* (Trondheim, Norway, August 30–September 02, 2005). Very Large Data Bases. VLDB Endowment, pp. 1128–1139.
9. Jeffery, S., Alonso, G., Franklin, M., Hong, W., and Widom, J. 2006. A pipelined framework for online cleaning of sensor data streams, in *ICDE 06: Proceedings of the 22nd International Conference on Data Engineering (ICDE06).* Washington, DC: IEEE Computer Society, pp. 140.
10. Jain, J. and Das, S. 2006. Collision avoidance in a dense RFID network, in *WiNTECH 06: Proceedings of the 1st International Workshop on Wireless Network Testbeds, Experimental Evaluation and Charaterization.* New York: ACM Press, pp. 49–56.
11. Levinson, M. 2003. The RFID imperative, CIO Magazine, December 1, 2003. http://www.cio.com/article/32004/SuccessfulUseofRFIDRequirestheRight Infrastructure (accessed October 10, 2007).

12. Niederman, F., Mathieu, R., Morley, R., and Kwon, IK-Whan. July 2007. Examining RFID applications in supply chain management, *Communications of the ACM*, 50(7), 92–101.
13. Owens, J., Chalasani, S., and Sounderpandian, J. 2005. Use of RFID in supply chain data processing, in *Encyclopedia of Data Warehousing and Mining*, John Wang (Ed.). Idea Group, Information Science Publishing, pp. 1160–1165.
14. Wagner, M., Clark, J., and Thomas, C. RFID industry survey: Measuring RFID use and performance in the DoD supply chain, March 2008, www.xiostrategies.com (accessed April 20, 2008).

第2部分 WSN

第8章 无线传感器网络中的地理位置路由

8.1 介绍

无线传感器网络由一套称为传感器节点的独立装备组合而成，拥有短距离的无线接口和用于监控环境参数（例如湿度、压力和温度）的硬件。传感器节点同样装有一个小型的微处理器，它们通常情况下是由电池供电。传感器利用其他节点传送数据到不在其覆盖范围内的其他传感器或者数据接收器。因此，对于无线传感器网络的分布式操作使用有效的路由协议是至关重要的。

无线传感器网络（WSN）从某种程度上来说和传统的 Ad hoc 网络相似，因为两者都是由不确定的和数量可变的节点组合而成，并都使用无线接口进行内部通信。因此，在 Ad hoc 无线网络中已经定义和测试过的路由协议同样可以使用到 WSN 的应用拓扑中。然而，节点的拓扑结构密度和拓扑结构频繁的改变是 WSN 的一个特有属性；不稳定性、能量稀缺、计算能力和节点的内存限制，阻碍了将传统的网络路由协议应用到正常工作的 WSN。

地理位置路由是一种依赖地理位置信息的路由技术。最早是由 Finn[1] 和 Takagi、Kleinrock[2] 在 20 世纪 80 年代提出来的，当时提出主要是为了无线网络。地理位置路由（也叫做 georouting 或者基于位置的路由）就是基于发送信息到目的地的地理位置而不是使用其网络地址的思想。在地理位置路由中，节点单纯地根据本地信息进行路由选择。节点仅仅需要知道自身的位置，目的地的位置，以及相邻节点的位置。有了这些信息之后，信息能够在不知道整个网络拓扑结构或者预先的路由发现机制下到达目的地。

缺乏全局信息意味着地理位置路由不能提供全局最优解决方法。因此，近似最优被广泛地应用于文献之中。事实上，这些协议局部操作产生了两个普遍的问题：路由回路和死头（dead end）。有一些不同的对策来处理这些问题，尤其是在单播的情况下。它们中的大部分是基于或者说是扩展了两个著名的技术：贪心转发和表面路由。贪心转发试图通过逐级逼近目的地来避免形成路由回路。如此一来，每一

个节点将信息传递到最合适的邻居节点（通常是从局部的观点来看）。

然而，贪心转发可能导致死头，即节点没有比其自身更合适的相邻节点。在这种情况下表面路由能够通过发现到达另外一个节点的路径来避免这种情况的发生，从而继续进行贪心转发。死头通常是位于未被任何传感器覆盖的区域的边界上。这些空区域也叫做空洞区。基本上，表面路由（也叫做周边路由）是用于包围空间区。

多播甚至比单播更加复杂，迄今为止只是提出了很少部分的地理多播协议。这里，问题是如何构建一条有效的通信树来到达一系列节点。WSN 的典型特征是在网络中部署了大量的节点，网络中的每个节点有几个不同的单跳邻居节点可以互相传递信息。因此，转发的替代选择呈指数方式增加。此外，多播树的形状直接决定了它的效率并从而决定了传感器节点的稀缺资源的使用情况。

此外，一些近期的研究工作提出了一项新的地理位置路由协议，叫做 Beacon-less 地理位置路由协议。这样做是为了消除在大部分的学术研究中一个共同的假设；即在单跳的邻居节点中使用周期性的短 HELLO 消息作为邻居发现机制。近期的研究结果表明这些信号在路由协议中引起很多的问题，例如干扰和碰撞。另外，它被认为是浪费带宽和能量的一种根源，尤其是在那些没有参加路由任务的节点中。

最后，需要重点指出的就是地理位置路由协议并不是完全的没有任何问题。某些方面还没有得到清楚的解决，例如节点确定目的地坐标的方式（定位问题）。还有一些可以测量的问题，如与传感器节点可以管理的有限规模信息相关的多播协议。而且，使用的典型 WSN 的单位圆盘图（UDG）会导致所设计的协议不能够很好地适应实际场景。换言之，覆盖范围不能假定为一个常量因为它取决于实际环境条件。这些仅仅是到目前为止，设计好的协议在实际场景中可以进行试验的问题的例子。

本章阐述了到目前为止在 WSN 中地理位置路由协议的不同方法。从单播场景开始，我们概述了在研究中最重要的协议，以及它们可能存在的问题。我们同样讨论了多播情况，然后提出了文献中的最新的 Beacon-less 地理工作。最后，本文总结了在以后的几年内地理位置路由协议的设计者所要面对的重要问题。

8.2 地理位置路由的原理

8.2.1 简介

由表驱动路由协议的性能和可扩展性问题激起，Finn[1]定义了我们现在称为地理位置路由的基本定理。Finn 过去一直在寻找一种方法来降低为了处理大的网络路由协议引入的不可接受的负担。另外，他注意到路由协议不能处理移动节点，另

外，他还想要降低管理路由表所需要的相当大的内存和计算能量。

Finn的想法简单有效，他提出在直角坐标系统中关联一个独特的位置到网络中的每个节点来替代使用地址去识别网络中的节点。考虑到这一情况，路由的任务减少到选择下一个消息将要传递的路由器。这样就创建了到达目的地的多条路径。已知目的地节点和其相关位置，每个路由器都可以确定与它相连的路由器中的与目的地最近的节点。然后，将信息转发到选定的路由器，使其越来越接近于目的地。

如今，大多数的路由难题来自于无线网络世界，在这种完全分散和非结构化拓扑中处理非均匀的移动节点。其中，WSN是一个特殊情况。WSN由大量的节点组成。另外，传感器节点在内存、能量和计算能力方面资源有限。

针对多跳网络提出的大多数路由协议是基于Bellman和Ford[3,4]所做的工作。一些著名的例子如自组织按需距离向量（Ad hoc On-Demand Distance Vector，AODV）[5]和最优链路状态路由（Optimized Link State Routing，OLSR）[6]。这些协议代表了两种不同的路由模式。AODV代表反应的路由协议设置（也叫做按需协议）。这些协议将路由发现推迟到需要它们的时候。另外一方面，OLSR代表了一系列主动或表驱动的协议，随着时间一直更新路由信息。更新过程独立于所需要的信息。

近年来，一些研究已经表明被动的和表驱动的路由协议，加上过多现有的混合方法，不能够应对无线网络提出的高要求，尤其是无线传感器网络。因此，随着Finn的工作应用于无线网络中，地理位置路由解决方案获得了很大的动力。

和传统的距离向量或者基于链路状态的路由协议不同，地理位置路由算法在单跳邻居节点上不需要互换路由表。节点仅仅根据自身的位置、目的地的位置、单跳邻居节点的位置来进行路由决策。通过几何条件来做出决定，所以节点的位置可以定义在一个公共的坐标系内。通常，坐标系定义为标准的R^2笛卡儿平面，但是也可以使用任何的能够计算几何距离的坐标系统。

无线通信的到来改变了设计网络和路由协议的方式。与有线网络不同，路由器使用无线接口能够与每个足够近的可达的设备通信。这取决于无线接口的覆盖范围和路由器之间距离的关系。显然，环境条件也能起到决定性的作用。因此，Finn的构想是在无线网络中利用位置来识别网络节点是更合理的，因为在这里位置和节点之间的距离对于确定连接是至关重要的。

而且，使用这种方法来确定路由就没有必要保存路由表了。因此，这个内在的，无状态特性对于构建大规模数量节点的路由协议是理想的。另外，因为使用地理位置操作来做路由决策计算能量需求非常低。基于位置算法的这两个特性使得它们非常适合运行于传感器节点上。

最后，无线通信采用共享媒介能够帮助减少带宽的使用，尤其是当多个节点需要接收相同的信息时。这种效应被称作为无线多跳优势而且能被看做为自由形式的广播。位于无线发送器覆盖范围内的所有设备接收信息，因此，一次的传输能够到达超过一个节点。这对于我们下面将要看到多播通信是非常有用的。

8.2.2　地理位置路由操作

自从采用了 Finn 的笛卡儿路由思想之后，提出了几个地理位置路由协议。每一个提议都提出了一个新概念、一个思想，或者针对一个具体问题的解决方案，但是基本上所有算法具有相同的基本结构。这个共同的结构大部分直接影响到地理位置路由的性能和可扩展性。而且，记住这算法的工作方式能够更好地理解单播通信与多播通信之间的区别。

通常情况下，在无线传感器网络中部署的所有节点采用相同的路由协议。源节点能够定期的发送信息，或者当外部事件发生时，中间节点仅仅在接收和数据处理之后才会传递信息。那些参与路由任务但是却不是源或者目的节点的节点被称为中间节点。

我们可以将一个标准的地理位置路由算法划分为下面的四个步骤。

（1）决定目的节点的坐标　目的节点的坐标是由信息的源节点决定的。对于中继信息这些坐标包含在信息的头部。坐标在地理位置路由协议中扮演地址的角色。我们随后将看到决定目的节点坐标的是一个附加问题。有一些方法来解决这个问题，但是总体上没有一个方法是完全令人满意的。

（2）决定一跳的邻居节点坐标　一跳的邻居节点坐标通常是通过周期性互换叫做信号灯的短消息来获取的。信号灯包含发送者的标识符和它的坐标。信号灯也允许节点知道它们的邻居，因为 WSN 拓扑不是静态的。这样做的理由是尽管传感器节点不经常改变它们的位置，但为了节省电池能量它们通常在一个责任周期内工作。即一些时间段内传感器是完全运作的，但是其他时间处于休眠状态。

（3）决定下一个中继　使用当前节点的位置信息、目的节点的坐标和一跳邻居节点的坐标来决定下一个中继。负责这一重要决定的功能通常被称为下一跳中继选择功能。大多数地理位置路由协议仅仅是对这个功能的定义不同。

（4）信息传递　在选择了哪个一跳邻居节点作为下一个中继之后，最后一步是传递信息。通常情况下，信息首部包括路由任务，即目的节点和所选择的下一跳中继节点之间的关联。所有的一跳邻居节点接收相同的信息，但是仅仅只有一个在路由任务中指示的节点继续进行路由。显然，多播情况下会有所不同，因为一些含有树分支的节点必须要传递相同的信息到不止一个邻居节点。在这种情况下，信息的首部能够包括一些路由任务。

这四个步骤基本上足够包括在文章中提到的不同规则的地理位置路由协议。另外，我们能够分析大部分作者设计地理位置路由协议时的一些隐含的设想。最重要的是假设源节点拥有决定其自身的位置和目的节点的位置（或者终节点）的能力。这些都是在实际中比路由本身更大的问题。

地理位置路由协议是基于节点位置坐标的，但是大多数方案不处理节点如何确定自身位置的问题。正常情况下假设传感器节点在某种程度上像 GPS 设备，但是

不可能在所有的情况下都这样。直到最近，GPS设备经常会消耗大量能量，而且过大的能耗导致不能够装备到由电池供电的典型的传感器节点上。现在，存在一些低能耗和小尺寸的GPS设备，但是GPS技术在室内情况下不能正常工作。

另一方面，发现目的节点位置的方法是多种多样的。我们能够发现的解决策略，从经典的集中模式，类似于域名系统（DNS），到非常复杂的负责记录特定区域内基于分布式散列表的节点的位置的模式。每个方案都有优点和缺点，但是它们都隐含了一些不希望的能耗。这一章专门讨论路由，因此，我们并不涵盖所有这些问题。所以，我们同样假定在文章中目的节点的坐标是已知的。

另外，确定一跳邻居节点坐标的过程也是一个关键因素，但通常被大部分协议所忽略。即作者假设网络中的所有节点有一个永久更新的表，包括了标识和所有的一跳邻居节点的位置。这通常是通过周期性互换短信息来获得的，也包含了必要的信息，然而，最近研究[38]显示，即使是在一个很短的长度，这些信号灯能够对路由协议的整体性能起到负面影响。事实上，由信号灯产生的干扰和碰撞会显著的降低网络吞吐量。在文章中这些问题通常不提及，因为在模拟阶段通常不考虑信号灯。然而，当测试真实的部署时这种效应是很明显的。

为了解决这些问题提出的一些建议，通常叫做Beacon-less地理位置路由协议。在这些协议中，一跳邻居节点的坐标是被动确定的。不采用主动的交换信息，节点使用一种有限一跳的传播形式进行信息搜索，寻找其当前的邻居节点。到目前为止取得的结果是非常可喜的。一些在真实场景中的实验已经表明Beacon-less协议在干扰和碰撞方面比标准地理位置路由协议能够更好地适应现实的环境。这对于无线通信是固有的。

8.3　地理位置单播路由

固定的坐标和单一的目标（单播目的地）意味着来自所有可能的节点位置确实存在信息应该遵循的清楚的方向概念。因此，每一个节点能够简单地决定从自身到目的节点的最佳多跳路径。节点和目的节点的距离假设为 d，r 为最大的通信范围。最佳的路径是到目的节点的一条仅含有 d/r 跳的直线。邻居节点恰好是最好的可能路径应该发起的位置的这种可能性是非常小的，但是知道了这个理想的位置就可能确定哪一个邻居节点会是最好的下一个转发节点。

然而，选择下一个节点的函数仅仅是依靠地理位置信息，因此，在最好的情况下，决策应该是本地最优的，但是不能保证全局最优。这就是为什么在本章中不同的协议使用不同的函数，使用不同的启发式方法去评估每个可能的下一跳候选节点的好坏。

另外，在某些情况下是不能遵循理想的路径。事实上，沿路径到达一个节点可能没有比自身更靠近目的节点的近邻节点。这些情况称为局部极小，必须采用其他

方式管理。因此，地理位置路由算法使用两种不同的操作模式：贪心的和周边的。无论何时尽可能使用贪心模式，周期模式严格局限于局部极小的情况下。

8.3.1 贪心方案

最初的地理位置路由的一个协议是 Takagi 和 Kleinrock[2] 的半径内最大前进（Most Forward within Radius，MFR）。通过定义进度的概念来确定最佳的候选者成为下一个中继转发者。进程的定义如下：给定的 s 为节点当前的持有的信息，d 为目的节点，n 是一个 1 跳邻居，n 的进度为定义为段 $\overline{sd}$ 和 $\overline{nd}$ 的长度的差异（见图 8-1）。MFR 的下一跳中继选择功能用来计算每个邻居的进度以及选择那个进度最大的节点作为下一跳中继节点。

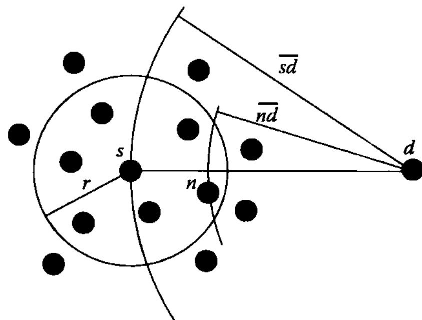


图 8-1　在 MFR 中，1 跳邻居被选作下一跳中继，提供最好的渐近。邻居 n 有一个 $\overline{sd}-\overline{nd}$ 的渐近

随后，由 Finn[1] 提出的笛卡儿路由是随机路由协议[9] 的一个改进版本，提出了一个简单的下一中继的选择功能。消息节点选择离目的地最近的邻居节点作为下一跳中继。这个算法是为由很多个网络节点组成的位于一般规则的几何结构中的无线传感网络设计的。然而，当我们在无线网络试着使用这种协议时，能够容易地找出协议进入循环的例子。例如，两个和目的地距离相同的节点将无限期地交换信息，因为每个节点都是另外一个节点更靠近目的地的邻居。

由 Kranakis 等人[7] 提出的指南针路由使用了一个不同的标准。这里，节点考虑自身和邻居节点片段之间的斜率。被选作下一跳中继的节点是那个斜率更靠近于从当前节点到目的节点的直线的斜率的节点。正如 Stojmenovic 和 Lin 在 Ref. [8] 中提出的，这个协议也能够导致失败。

8.3.2 周边方案

如前所述，地理位置路由协议需要建立机制以恢复经常被称作为局部极小和无效地区的那些情况。有些协议，例如笛卡儿路由协议，提出当被路由的信息到达了一个死路时使用受限洪泛。然而，这个不能够被认为是一个纯地理位置路由技术。另外，形成复本信息的可能性，以及引入的代价，使得这些解决方案不适用于无线传感网络。

大多数作者用图来模拟无线传感器网络，顶点代表传感器节点，边线代表了节点之间存在的直接通信。这是非常方便的，因为有一些来自于图形理论的非常有名的算法对于解决特定的路由问题是非常有用的，例如 Dijkstra 在图中的寻找最短路

径的算法。

右手规则[13]是一个能够找出迷宫出口的算法。算法的工作原理如下。游戏者行走时保持一只手接触迷宫的一面墙，例如他的右手。如果迷宫不包含分离的部分，那么只要有出口游戏者一定能够到达。在一个无线传感器网络中，一个空洞区域可以被看做是一个迷宫中的一个房间，它的墙被图片的边所定义来模拟这个网络。在空洞处的一个顶点出发，可以应用左手法则或者右手法则（见图8-2）来包围它。

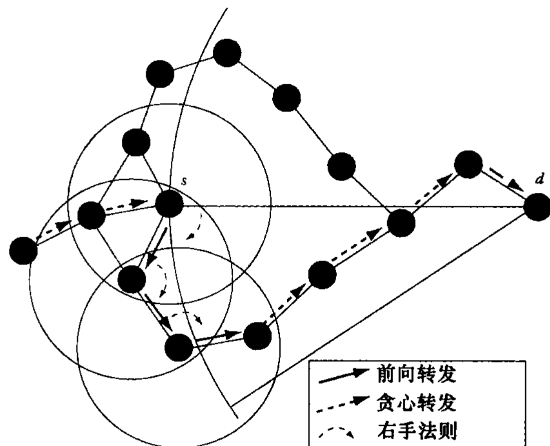


图8-2　在利用前向转发包围一个空白区域中应用右手法则的例子
（朝d转发，节点s是一个本地最靠近d的点；因此s发起参数转发，
直到它搜索到一个比自身更靠近d的节点，然后继续进行贪心转发）

这个算法的唯一要求就是运算中不能出现交叉边缘，即这个图必须是平面的。定位操作作为地理位置路由算法的关键之一，我们不能假设节点知道整个网络拓扑结构。每一个节点仅仅知道整个图形的局部状态，即本地子图。

幸运的是，有一些拓扑控制算法能够被用来使得每个节点获得平面视图图形的一部分。这些平面算法能够被应用于本地[14]。具体地，这些算法仅仅使用那些嵌入了当前节点和单跳邻居节点的子图来做运算。

最常用的平面算法是加布里埃尔图（Gabriel graph）[12]和相对邻域图（Relative neighborhood graph）[11]。它们通过删除连接当前节点和它的邻居节点之间的边来工作。还有一些其他的算法，其平面子图有细微差别，但是由于具有较高的计算代价而较少的使用，例如，局部德洛内三角算法（Localized delauney triangulation）[15-17]和莫雷利亚测试（Morelia test）[18,19]。

针对处理空洞区域有不同的方法，其中的大多数是基于在局部平面子图上使用右手规则的思想。在它们之中我们评论了由Bose等人[10]提出来的FACE-2，随同FACE-1（一个更加简单和低效率的最初版本）。在FACE-2中，信息从源节点到目

的节点的路径通过内部的邻近的表面（由节点之间的边限定的封闭的多边形区域）通过一条位于源节点和目的地之间虚构的线进行分割。例如，根据右手规则出发，信息第一次通过由假想的线分割开的边缘；信息继续应用相反的规则，在这种情况下是左手规则。因此，信息每次从一个面到下一个面，规则随之改变。FACE-1 和 FACE-2 的性能都很差，因为它们建立的路径可能太长。因此，这些作者提出第三个联合贪心路由算法和 FACE-2 周边路由协议的算法。新算法叫做贪心对贪心（Greedy-Face-Greedy，GFG）[10,14]，因此是个贪心的方法，它通过使用 FACE-2 算法能够解决空洞区域的问题。因为周边路由（FACE-2）仅仅当需要时才会使用，由 GFG 创造的路径比那些由 FACE-2 建立的要短。

8.3.3　处理真实情景

文章中所提到的大多数地理位置路由协议，是在理想环境下设计的。例如，假设一个理想媒体访问控制（MAC）层或者考虑网络作为一个 UDG。假设一个理想的媒体访问控制层意味着协议不是设计用来处理无线连接固有的损耗和传输错误的。另外，真正的无线链接没有一个固定的通信范围，因为假设的是网络使用 UDG 来模拟。

比如 Zhao 和 Govindan[20] 提出来的一些研究，以及由 Woo[21] 所做的工作，确信了无线链接和理想中的媒体访问控制有很大的差距。因此，考虑到在实现实际的解决方案路由过程中消息的损耗和干扰是肯定会存在的。

另外，两个节点仅仅只有当它们的欧氏距离小于或等于常数时它们会被一条边相连，UDG 是一个图。这个常数与一个节点的最大无线通信半径有关。因此，两个节点之间直接的通信仅仅由它们的位置决定，不考虑由环境因素引起的无线信号的改变。

此外，近期的研究表明不仅仅是距离和环境，而且信息的大小同样对信息的传输率是很重要的。具体来说，由 Sánchez 等人[31] 研究的工作清楚地表明数据包的大小和数据包接受率（PRR）之间有明显的关联，在某种程度上，数据包越长，数据包的接受率越低。

显然，这些假设降低了设计协议的效力，甚至可能降低模拟运行所需要的时间。然而，该协议是在理想条件下设计的，不能期望在实际工作中正确地工作。在无线传感网络中设计路由协议时能量消耗被认为是另外一个重要的因素，传感器节点通常是由电池来驱动的，电池有一定的使用寿命，而且在大多数实际情况下部署新的传感器比更换电池要容易得多。这对于无线传感网络来说是尤其关键的。因此，减少能量的消耗付出了很大的精力。

能量有效的地理位置路由的局部最优源路由（LOSR）是这个领域最新的方案之一。由于 LOSR 只影响协议的信息传递，因此它能够帮助每一个地理位置路由协议减少能量消耗。这个想法是按照最短路径的思想，考虑到能量，通过应用 LOSR

的地理位置路由算法来寻找当前结点与最初被选作下一跳中继的节点之间的最短路径。

通过使用著名的 Dijkstra's 算法能够计算图表中两个节点之间的最短路径，但是需要知道整个图表的情况。LOSR 只在当前节点的本地子目录中使用 Dijkstra's 算法做出路径决定。即子图表仅仅是由此节点和其单跳邻居节点组成的。本地子图的边缘标有把一个消息成功地传递到链接的另一端所需要的估计的能量消耗。然后，通过应用 Dijkstra's 算法我们能够计算出最低能耗路径。再者，LOSR 占用能量可能是由于在处理无线链路的性质错误时有消息重传。

一旦计算出来最短的能量路径，消息通过这条路径发送。要做到这一点，LOSR 使用一个包括必须穿过的节点列表的源路由头（SRH）。根据当前节点的局部信息，使得信息通过最佳能效路径传输（最终通过那些不提供一个改进反而降低能量损耗的节点）。另外，因为源路由头经常引导至提供优势的节点，即最初选作下一跳的节点，从而避免路由循环。

8.4 地理位置多播路由

无线传感器的应用是十分广泛的，包括监控、分布式控制等。在某些情况下，传感器网络也与环境相互作用。例如，一个监测站能够发送命令到不同的传感器节点，控制这水龙头打开还是关闭水阀。在很多的实际情况下，一对多和多对一通信成为整个系统的基本构建块。在这篇文章中，高效的多跳路由协议能够传输数据或者命令到预定的目的地（传感器节点的一个子集），且相对于广播或者多播通信能够提供扩展的带宽和节省能源。

此外，想象一下消防队员进入一栋有火险的建筑物内。每一个消防员都希望接收到在建筑物内部的传感器所发出的警报信息。探测到火警信息的传感器可能在离消防员很远的位置，而且消防队员也可以分布在建筑物内。因此，建立一个可以对全体消防员传递警告信息的可靠机制是重要的。

即使在知道整个网络拓扑结构的情况下建立高效的多跳树也是非常困难的。对于固定网络，最小的多跳树叫做 Steiner 最小树（SMT），其计算是一个完全 NP 问题[32]。对于无线多跳网络，SMT 不是最优解，计算最佳树也是完全 NP 问题[33]。因此，在这篇文章中计算有效多跳树的大多数方案是采用启发式的解决方案。尽管它们没有提供最佳的解决方案，但是它们能够在能量消耗方面，以及信息开销和计算成本方面有很好的表现。

在文章中移动 Ad hoc 网络（Mobile Ad hoc network，MANET）的多跳问题已经被广泛讨论。在过多的现有解决方案中，我们能够引用一些非常有名的方案，例如 MAODV[34]、ODMRP[35]和 MMARP[36]。但是，已经证明这些方案不适用于无线传感器网络[37]，因为无线传感器网络不同于传统的 MANET、尽管它们之间存在一

些相似之处，但是分歧明显多于相似。例如，组成 WSN 的节点的数量通常比 MANET 的节点数量要多很多倍。而且，在能量、内存和计算能力方面，传感器节点通常比 MANET 节点（如笔记本，或者甚至像掌上电脑之类的设备）拥有更少的资源。

对于地理位置多跳路由（GMR）的特例，这篇文章中大多数协议是基于 GFG[10]的扩展，因为它是最著名的地理位置单播路由算法之一。然而，将一个单播地理位置路由协议应用到一个多播情况下不是一个简单的任务。下面一节将具体介绍这种应用是怎样执行的。

8.4.1 从单播到多播

给定一个单播地理位置路由协议，将它用于处理多播目的地，更简单的方式是分别传输相同的信息到每个不同的目的地。这种更改（也叫做多播）在执行性能上来看不是一个好的解决策略。显然，信息将被相同的节点传输很多次，具体情况是，那些节点通过两条或更多的路径在源节点和目的地之间是共享的。但这种最原始的解决方案毕竟执行能力很低，也同样说明一些作者将一个地理位置单播路由更改到多播情况下所作的努力。

改变一个地理位置单播路由算法到多播状态需要改变路由过程的各个方面。不像在单播情况下仅仅只有一个目的地，我们只要去寻找单一的路径，我们应该考虑到，在多播情况下，目标是（需要）建立一棵连接源节点和每个目的地的树。得到的路径树将会包括不同的开始分支的节点和相同的继续分支的节点。每个分支延伸到目的地的子集。因此，确定了哪里开始分支决定了树的全部路径。这样也直接影响到了传输的效率、具体的带宽和传递一个信息到所有的目的地所需的能量。

虽然单播地理位置路由算法的基本操作依然是有效的，但是在 GMR 情况下，由于大量原因问题变得更加的有挑战性。首先，节点需要知道每个信息的所有目的地的位置。因此，信息中必须包括了每个目的地的位置，源节点必须确定这些位置。

和在单跳情况下的相同，在多跳情况下，每个信息必须包括所选的下一跳中继，以及它所负责的接收者。因此，在每个节点收到信息时检查自己现在是否是中继，然后检查是否是目的地节点之一，最后，运行路由协议来决定对信息该做什么。在这种情况下，有两种选择：将信息仅传递到一个中继或者更多的中继，从而建立一个分支点。如何建立分支，是最重要的挑战之一。

然而，当节点决定把路径划分为两个或者更多的分支时，中继的选择包括决定路由到达目的地子集应该考虑哪个下一跳。因为可能的选择数随着邻居和接受者数量的增加而呈指数增长，传感器节点需要一些启发式方法来运行这些协议。图 8-3 表明当一个数据包的路由目标的集包括五个节点时，有三个邻居的节点 s 的不同的可能性。另外，选择一个或者多个下一跳（也就是说，是否去创建一个分支）的

决定对于整个多播树结构的成本有着重要的影响。

在传递阶段，单播和多播的不同也是显著的。在选择了最优的下一跳中继候选之后，当前作为一中继的节点接收的数据包一定要被传输。总体上，大多数现存的方案已经假设在理想的通信信道中。然而，在实际情况下数据包可能丢失或者损坏。协议应该可靠地传递信息到所有已经选择的下一跳。否则，一个单一的数据包丢失可能引起多个目的地丢失数据包。

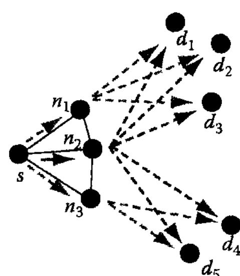


图 8-3 随着邻居和目的地的增加，多播路由潜在地增加了很多替代方案。这个例子显示了对于有三个1跳邻居的节点 s 向五个目的路由的两个可选择的方案

另外，当一个节点能够自身决定开始一个分支，信息头部必须不仅包含一个单一的中继和目的地对，而且要含有很多新创建的分支。最后，将周边模式用于路由到多个目的地，这些目的地需要仔细设计避免造成路由环，或者用周边模式来处理在贪心模式下的一些可路由目的地的情况，和在周边模式下的其他一些情况。

8.4.2 多播贪心路由

基于位置的多播（PBM）路由协议，由 Mauve 等人在参考文献［23］中首次被提出的。这个协议，尽管最初不是针对传感器网络提出的，但是符合本地标准和无线传感器网络运行路由协议的有限的网络代价需求。PBM 是 GFG[10] 的泛化，路由到多个目的地进行操作。它建立了一个多播树，树的形状根据一个标记为 λ 参数从最短路径树到一个合适的最小代价多播树变化。

正如我们前面提到的，GMR 协议最重要的一个方面是选择转发的邻居。协议的这一步操作是通过控制参数 λ 来实现的。作者试着在传递信息总的节点数和到达目的地最优的个别路径之间找出一个好的均衡。

为了选择邻居节点，每一个节点使用下面的函数评估所有可能的邻居子集来预计最优的替代选择。

$$f(W)=\lambda N+(1-\lambda)D,\ 0\leqslant\lambda\leqslant1$$

式中 N——被邻居总数量划分的被选择邻居（$|W|$）的数量；

D——W 中的节点到目的地的最小距离的总合，用从当前节点到所有的目的地的距离总和来对它进行标准化。

当 $\lambda=1$ 时，选择的邻居的数量是最小的。因此，多播树分支尽可能晚的建立成为近似于 Steiner 树。当 $\lambda=0$ 时，目标函数奖励被选择的节点提供最短的路由到每一个目的地的替代节点。在这种情况下，得到的多播树类似于最短路径树。

从所有可能的邻居的子集（W），当前节点选择最优的 $f(W)$。如果最佳的邻居子集是一个单一的节点，则那个节点对于所有的目的地将会是唯一的中继。如果子

集包含超过一个邻居，子集中的每一个节点将传递路由数据信息到一些目的地。如果在一些节点处没有节点提供朝向一个或者更多的目的地的中继，那么发起一个不同的分支，个别情况下，对于那些目的地应用 GFG。

这个方法的主要问题是决定最优值 λ 不是一个简单的过程。事实上，作者计算了不同的 λ 值，但是从来没有找到一个最优值。另外一个问题是算法的计算消耗。每个节点的每一跳传递一个信息，节点必须测试所有可能邻居的子集，并在它们之间选择最优路径。因此，假定邻居为 n，则可能的子集是 $\sum_{k=0}^{n}\binom{n}{k} = 2^n$。这样就导致了一个指数级的时间复杂度，复杂度过高会使传感器节点计算能力不可行。对于有大量接收者的网络，由于多播数据包中需要包含所有目的地，PBM 可能表现不好。基于位置的可扩展的多播移动 Ad hoc 网络（SPBM）[24]被设计去提高可扩展性。它使用节点的地理位置提供可扩展的组成员规则和转发数据包。SPBM 主要关注于以可扩展的方式管理多播组。

然而，SPBM 不能够提供一个有效的组播转发，因为其对于每个目的地使用一个单一的地理位置路由，也就是，一个多播方法。再者，邻居之间路由表的积极交互使得协议对于多播组的数量是不可扩展的，就像 PBM。此外，保持从属关系表意味着节点需要在内存上保存大量的信息，这和地理位置路由的原理是不一致的。

GMR[25]整合了一些 PBM 可扩展性问题的解决方案，同时在带宽的使用方面获得更好的结果。事实上，由 GMR 计算的多播树的效率和当网络的密度足够大时最好的集中启发式算法效率一样好。GMR 是基于成本比进度框架的无参数算法。具体地说，为了确定目的地节点最好的分割，对于每一个子集必须选择传递中继的邻居节点，在 GMR 中，节点使用替换代价和它产生的进度之间的比例评估不同的选择方案。

考虑多播树总体上的形状，分支太早和分支太晚都不好。成本与进度比试图在这两个极端情况下找到平衡，因此当使用这些已选择的邻居作为下一个中继节点，成本可以用被选择的邻居超过 1 的个数来表示，这不利于分支和计算作为到达目的地的前进距离的进度。

结果表明在一系列的实际网络中不考虑 λ 的值，GMR 比 PBM 要好得多。主要的原因是 GMR 的邻居选择功能能够在一个多播树中获得非常好的近似最小带宽消耗。此外，在相同的工作中，由于采用了贪心集合并算法使得 GMR 在计算时间方面更加高效。利用该算法实现了对最佳分支选项的近似多项时间复杂度。

GMR 是一个对于减少带宽使用最优的多播地理位置路由协议。其主要的目标是使用尽可能少的节点建立多播树，因此传递一个消息到目的地所要传递的次数较低。减少传递的次数同样能够被看做为减少多播网络传输信息所需的能量。虽然如此，相同的作者提出了一个新的叫做局部能量有效多播算法（LEMA）的协议，如参考文献［42］所述，这个协议是专门为建立尽可能低能量消耗的多播树而设

计的。

在参考文献［43］中描述了LEMA和最近由Frey等人提出来的叫做最小生成树的能源感知组播路由（MSTEAM）。上述两者是基于多播骨干的使用，相比于GMR和PBM这个方法是完全不同的。特别是，节点采取了两种决策：关于分支的决策，即在大量生成目的地子集的分支中如何以及何时进行信息划分，决定每一个分支的下一步跳中继。

一个节点路由一条消息来计算由它自身和目的节点构成的图的最小生成树（MST）。然后这个树被作为一个向导而去确定何时和怎样发起分支。当当前节点和一系列目的地节点的最小生成树有许多发起于当前节点的边时，需要进行信息划分。这些边中的每个边生成的目的地节点集合到一起。在MST中直接和当前节点连在一起的目的地节点也和当前节点最近，被称为第一级目的地，因为伴随着MST的计算，信息在到达其他目的地之前应该先通过它们。

然后，当前节点选择提供一个更大的朝向每个一级接收者的前进的1跳邻居作为下一跳中继。下一跳中继将会负责路由信息到作为目的地之一的一级接收者，在MST中通过它生成其他的目的地。显然，相同的1跳邻居可能被选作两个或多个分支。

此后，当前持有信息的节点计算传递信息到被选作下一跳中继的一跳邻居子集的最可能的方式。为了达到上面的情况，在本地仅由自身和其邻居组成的图表中使用Dijkstra算法来得到最低能量路由树。Dijkstra算法寻找最短路径，因此只需使用在一条边上传递信息所需要的能量级别来标记本地图表的边界问题就能够解决。

最后，每一个在信息的头部的路由分配同样包含完整的最短路由计算，因此使用最佳本地路径来选择作为下一跳中继的一跳邻居。这可以被看做是一种本地源路由，但是通常只有少数中间跳。路由不包含节点第一级接收者的前进走向，但由于使用了源路由头这不是一个问题，即是，它是不可能在这个阶段进入环路。由LEMA获得的结果是很好的，因为在一些实际情况下与使用集中化算法能够获得的结果是非常接近的，例如MIP[41]。此外，对比发送的信息数，即所用的带宽，LEMA的结果比GMR仅仅差一点。

8.4.3　多播周边路由

在单播情况下，在贪心路由由于缺少邻居提供朝向目的地的前进而无法继续的位置，使用不同的基于GFG的本地协议来解决寻找一个节点的问题已经被广泛地接受。然而，同样的问题在多跳情况下也同样存在。一些协议求助于GFG，但是在很多情况下，对于相同的节点，在贪心路由方式下有超过一个目的地不能够达到。这些情况，在路由环路方式下多跳目的地使用先前的提议处理结果不好。总体上，采用多重单周边路由任务（每个目的地一个），这样浪费网络资源。尽管平面路由是由单播情况发展而来的，最近将它与多播协议结合起来处理多目的地的情况。一

些已提出的解决方案分别处理每一个目的地，因此导致能量消耗的增加。将平面恢复扩展到目前描述的多播情况可能会限制特定的平面图或者不能提供传递保证。

幸运的是，MSTEAM 的作者扩展了使用多播主干来设计一种新的多播平面恢复模式。正如我们已经讨论过的，多播骨干是至少包含当前持有的信息和它所负责管理的目的地节点的欧几里得生成树。这种思想是使用多播骨干作为向导试着按照骨干形状的边缘去传输一个多播信息到所有覆盖的目的地节点。同时，随着消息靠近目的地，多播骨干的形状与真正的网络拓扑相适应。

多播周边适应被称作为 MFACE[44]，当在贪心模式下路由时，对于无法访问的目的地子集它使用和多播骨干相似的思想。然后，任何起源于当前节点的持有信息的骨干边缘将生成一个新的消息副本。随即，每个副本朝着相应边上的目的地节点路由。每当消息到达一个平面边缘，这个平面边缘由不同于初始的平面边缘的骨干边缘所分割，消息便分成两个副本，处理一个由在交叉点划分多播骨干定义的多播目的地的分离子集。MFACE 的行为和深度第一树探测方法类似，但是在同一层有平行的探测子树。

8.5　信标减地理位置路由

8.5.1　动机

本地化的操作，降低了计算和存储要求，而且更重要的是，需要节点数量的可扩展性是地理位置路由算法最重要的特征。因此，在 WSN 领域路由协议使用这个技术变得非常的普遍[39,40]。正如我们已经提到的，大多数作者提出的假设是节点已知其单跳邻居节点的位置。事实上，使用著名的信标机制就可以很容易地得到。信标是包含发行人的标识符和其坐标的短消息。因此，假设节点知道其自身的位置，而节点发送此消息包含在一定的周期内的信标，每个节点都能够知道其单跳邻居节点的位置。

通过增加信标传输的频率，该机制对于应对网络改变的坚固性更好。拓扑改变通常由流动性的节点，新节点的加入，或一些节点寿命的终结引起，甚至由于工作周期限制大多数传感器节点采用减少能量消耗。然而，在整个网络中，更高的频率意味着更高的传输速率，我们必须指出网络中的所有节点做同样的处理；因此，每个信标循环代表 n 个信息，n 被认为是节点的数量，即它可以被看做是一种流。

从节约网络资源的角度来看，网络中流动速率高不是最好的做法。另一方面，减少信标的频率会导致节点对其邻居节点有一个错误的认知。该信标机制的频率是一个很难调的参数，且它甚至可能取决于实际情况。

另外，信标不仅能够产生大量的网络通信量，而且可能干扰正常的数据通信而导致数据丢失。显然，信标同样会从传感器节点上消耗能量，但是最重要的问题是

所有不想要的网络流量、干扰和能量消耗影响网络中的所有节点。换句话说，由于这些问题传感器节点没有参加任何的路由任务，不像节点路由信息，它们不会因为有一个精确的单跳邻居表而受益。而且，尽管采用高刷新率的信标机制，近期的研究显示了当考虑流动性时地理位置路由算法的局限性[38]。

为了避免这些定期的传输，信标减协议在文献中被提及。总的思路是当路由数据包需要选择下一个转发器时被动地发现邻居位置信息。信标减路由协议试着通过推迟发现一跳邻居到最后一步，来代替保持邻居表主动的更新，从而节省网络资源。

8.5.2 非协作方式

在最开始设计的信标减地理位置路由协议中，使用简单的询问回答选择机制能够主动地识别邻居，即持有数据信息的节点发送一个控制信息，表明它发现邻居节点的兴趣（见图8-4）。由于共享媒介，信息能够被所有一跳邻居接收到，因此每个接收节点都能够回答一个与信标相似的控制信息。反应包括标识符和邻居节点的位置。节点接受所有的回答后能够决定选择哪一个邻居作为下一跳中继。最后，数据信息被传送，包括所选择的下一跳中继的头部的标识符。

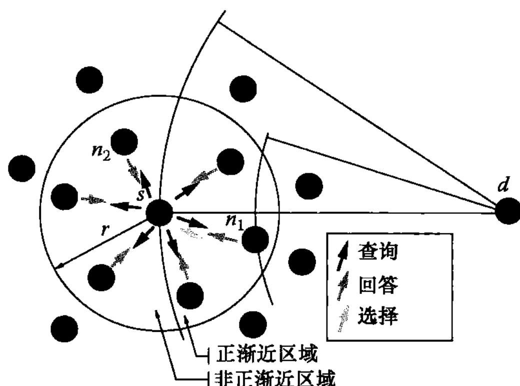


图8-4 在三阶段无信标路由的集中方式方法中的大体操作：查询、回答和选择。在这个例子中，s广播一个查询，然后一跳邻居回答表明它们的身份和位置，最后s选择位于正渐近区域中的最好的下一跳中继节点

这有一种简单的方法允许我们将任何标准的地理位置路由算法转换为一个信标减地理位置路由。然而，这个原始设计存在一些问题使它变得不切实际。首先，我们必须指出因为单个节点这个方案能够被分类成非协作的，当前持有信息的节点，负责路由决策。这本质上没什么不好的，但是我们在朝向目的地的路径的每一跳至少需要三条信息。假设一个典型的高密度无线传感器网络，每一个节点将拥有远远超过一个的邻居；因此，每一个单跳的信息量是非常高的。

还有，让每一个邻居在接收到查询的时候，应答可能不是最好的选择，因为由于碰撞丢失信息的概率可能非常高。另外，一些邻居节点的应答也是不必要的。根据贪心路由方法，所有位置比当前持有信息节点要远的邻居都不能够被选择为下一跳中继；因此，它们的应答是没有必要的。

基于争用的转发（CBF）[28]和隐式地理位置转发（IGF）[27]是第一个设计替代

方案的很好的例子。正如我们已经详细描述的，两者都使用相似的三阶段（RTS/CTS/ACK）握手过程来选择下一个转发。在 CBF 中，持有信息的当前节点传送一个准备发送（RTS）的控制命令，然后候选邻居节点全用一个清除发送命令（CTS）应答，最后，使用确认信息（ACK）执行选择下一跳转发。IGF 的操作和它非常相似，但是包含一个另外的控制信息。原因是当在网络层中使用两个信息执行选择下一个转发的时候，邻居节点的发现（头两条信息）在物理层上实现。第一条确定了选择的下一跳中继，第二条来自于被选择的邻居的应答，确认数据信息的接收。

这两个协议都纳入了相同的机制来避免由一跳邻居应答引起的碰撞。在回复之前强制邻居延迟一段时间。这种延迟是基于邻居节点之间的位置、持有信息的当前节点位置和目的地节点的位置之间的关系。因此，那些位置离目的地更远的邻居节点在回复之前延迟一段时间。其优点是双重的：第一，由于回复的产生不在同一时间段而减少了它们之间碰撞的次数，第二，强制最适合的邻居生成下一中继而率先回复。这样允许其余的邻居节点取消了回复，也减少了全部的推迟（也就是说，第一个回复收到后就立即开始转发）。

具体地，CBF 集成了一些附加特性来尽量保证在传输的每一步仅有一个回复。其思想是确保所有到下一跳中继的候选邻居节点能够听到传输的第一个邻居的回复。为了这个，作者使用转发域的理念，即仅仅在预设区域的邻居能够被看做是（考虑成为）下一跳中继的候选者。这样的话，作者以这种方式定义一个这样的区域，给定一个预先定义的覆盖半径，所有在其中的节点之间能够相互通信。

8.5.3　协作的方式

正如我们所看到的，非协作方式有一些缺点，尤其是对于每一个单跳的最小信息数是三这样一个事实。这样通过增加控制开销降低了协议的整体表现。因此，有些研究者通过使用完全不同的方法尽量去解决这些问题，即，在该分布式模式中，邻居之间互相协作决定下一跳转发节点。

显然，下一个转发的分布式选择不能够使用复杂的协议来保证选择的是一个单一的节点。在那种情况下，这个方案比集中式效果还要差因为它通常意味着几个阶段和多个控制信息。另外，邻居节点竞争成为下一跳中继。事实上，使用延迟响应的机制，正如我们上面描述的，赢得竞争的是第一个回复的邻居，因此，成为下一跳中继。

BLR 协议[26]充分体现了这个分布式操作。BLR 简化了选择下一个转发的过程，来减少每一步所需的信息量。和强制邻居回复 CTS 信息，然后决定下一跳中继的非协作方式不一样，在 BLR 中，邻居自行决定是否适合作为下一个转发。当邻居认为其是很好的候选者时，它直接发送消息。和 CBF 和 IGF 的定时器方式相同，最优的邻居应该最先回复；因此，其转发的信息作为事实上的下一跳中继。其余的

邻居期望听到转发，因此，取消其自身的定时器。

正如在CBF的情况下，那些定时期最先过期的节点的邻居的位置远离节点的覆盖范围，将得不到它的传输。因此，它们自身的定时器将最终期满，而使得其进行第二次传输而第二次传输是不必要的。为了解决这个问题，作者也建议在区域中限制邻居能够自行考虑下一跳中继候选者的区域，而避免同样在CBF情况下的重复。这种转发区已经定义，因此所有覆盖的节点能够彼此通信（见图8-5）。

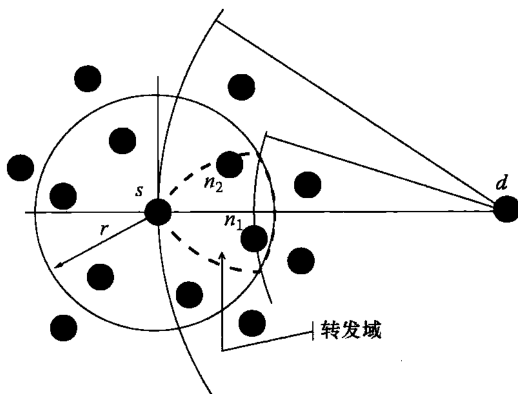


图8-5　转发域限定了邻居子集，邻居子集中的节点能够监听彼此的传输

8.5.4　处理空洞

到目前为止所描述的两种不同的信标减方案能够应付标价贪心转发，但是这些方案不能够应付那些节点没有提供一个积极的前进到目的地的情况。使用非协作方式，因为邻居都不在适当的区域内，没有能够回复原始的控制信息。分布式方案存在同样的问题，但另外，持有信息的当前节点不能够探测到它，因为它的使命是让其转发信息以及让邻居自己决定下一跳中继的邻居节点。

一些方案，例如CBF，借助一些知名的周边转发机制来处理这些情况，例如定义在贪心的路由[10]中的。由于工作情况下的这些机制在本地子图的平面版本中应用右手规则，有必要知道完整的邻居列表。因此，建议替代方案发送表明完整的邻居发现请求的一个特殊控制信息。也就是，所有的邻居节点必须独立回复其位置。这代表一个附加的控制信息开销，同时增加了由于碰撞的丢失信息的概率。

为了解决这些问题，Chawla等人[29]在回复阶段通过减少回复的次数来改进算法。他们提出一个新的计时器分配功能，如果一个交叉链路在它们添加到平面图中出现，避免邻居应答。也就是，更远的邻居第一个回答。在那种方式下，一个邻居能够决定是否其属于平面图。

另外，也有一些方案没有使用周边转发方案。例如，在盲地理位置路由(BGR)[30]中，我们能够找到一个原始替代。BGR是一个和CBF非常相似的信标减的地理位置路由算法；其主要贡献是它的恢复方案。和CBF不同，这里转发节点可以尝试三个不同的转发区域来寻找邻居的回复。协议尝试的第一个区域是一个覆盖可能的提供朝向目的地的可能的邻居的区域。假如没有收到回复，使用一个新的转发域。第二个区域来源于旋转第一区域向左或者向右60°。假如仍然没有接收到回复，使用最后的区域，从把最初的区域转换到相反的方向上。

8.5.5 处理实际场景

一些相关的协议是在假设信息不会丢失的情况下设计的。然而，在实际的情况下丢失一些信息是非常普遍的。因此，如转发区域的方案不能够保证来自最有可能的邻居节点的单一的回复。它的响应不能够到达所有的邻居，因此，他们将响应到定时器期满。在如 CBF 或者 IGF 的协议中，当在实际情况中测试时响应的次数提升，但是最坏的问题发生在遵守协作模式的协议中。例如，在 BLR 中，每一跳产生的重复路径的数量显著增加。这代表了控制开销的增加和到达目的地的重复信息的数量增加。

而且，节点使用的定位系统的准确度可能没有要求的精确。例如，因为仅仅使用坐标和距离来计算延迟，一些节点能够得出它们在同一个位置，因此同时响应。

另外，接受消息的可能性不仅受到距离的影响，而且受数据包大小的影响。因此，邻居节点能够因为之前接收到一个控制信息而被选择，例如上面所提到的发现请求，其响应也是第一次被接受。虽然如此，当数据信息被传输时，因为这些消息通常比那些控制信息要大，可能无法正确接收信息。

如参考文献［31］所述的 BOSS（传感器网络信标减按需策略）算法，在设计时考虑了无线通信典型的损耗和碰撞特征。BOSS 协议和 CBF 相似因为它也使用三阶段握手，但是处理真实场景的方式不同。例如，在 BOSS 中使用的发现请求包含完整的数据有效载荷。因此，只有接收到数据包的邻居节点才能参与下一阶段。这也保证了被选择的下一跳中继接收到了数据包，所以能够有效的执行中继。

再者，努力克服节点坐标不精确，BOSS 定义了一个新的时间分配功能叫做离散动态转发延迟（DDFD）。DDFD 根据朝向目的地的进度把邻居区域分成分区。位于相同分区的邻居共享相同的基本延时。计算最后的延迟增加一个随机的毫秒数。因此，在分区中有一个高进度的邻居能够在更远的分区的邻居之前回答，而且由于延迟的随机部分导致在相同的区域内的邻居不能够同时传递信息。DDFD 的目标是减少在选择阶段响应的碰撞数。

最后，BOSS 拥有另外一个特性来减少形成重复的可能性。邻居听到来自不同节点的回复响应后取消自身的定时器，但是当选择信息到达时同样取消定时器。即持有信息的当前节点仅仅当其接收到第一个响应后传递选择的信息，因此，那些因为没有接收到任何的响应而没有取消他们定时器的邻居节点，能够立即这样做。

8.6 总结和讨论

地理位置路由算法代表了对于开发 WSN 应用的一个基本构建块。很明显这些协议的整体性能和可扩展性优于先前的方案。然而，我们必须认识到还有很多的工

作要去做。有一些问题能够找到更好的解决方法，而且有些问题尚未解决。

正如我们所看到的，对一个具体节点位置的确定是一个路由协议的先前阶段，这个问题甚至比路由本身的问题还要大。到目前为止提出的方案已经被接受，但是所引入的开销仍然很大。对于路由自身，对空洞问题（也就是说，GFG和它的变种）最普遍使用的解决方案仍然未在真实环境中进行完全的测试，但是很容易看到的是这些策略过于依赖高度精确的位置信息，因此当处于真实环境中很可能失效。

当今在传感器节点中的硬件允许路由协议控制传输能量。这能够用于减少能量消耗，但是功率调节可能导致覆盖范围的变化，而更加重要的是，数据包接受率的变化。大多数设计用来降低能量消耗的协议是基于能量模型的，在这个模型中，能量消耗是和距离相关的。然而，调整传输功率只到达期望的特定的位置的节点可能改变这个节点接受消息的概率。因此，考虑新的能量模型，至少，像环境、距离、传输功率和数据包大小这些因素都应当被考虑进去。

多播协议还有附加的问题。具体地说，我们必须强调与目的地数相关的可扩展性问题。因为信息的头部包含了每一个目的地的位置，很明显这个数目是有限的，尤其是在WSN中，信息的大小受限。

最后，尽管信标减协议解决了一些由信标引起的和网络开销相关的问题，也有一些悬而未决的问题影响到它们。从我们的角度来看，最重要的是大多数信标减机制对于一个固定的最大覆盖范围有强烈的依赖性。大多数定时器和转发区域决策是基于覆盖范围所涉及的距离。因此，在真实的实验台中，如果一个远离预配置的覆盖范围的节点是可达的，或者相反地，位于覆盖范围内部的节点由于传输错误而不可达，我们能够发现一些古怪的现象和错误的决定。

参考文献

1. G.G. Finn, Routing and addressing problems in large metropolitan-scale internetworks, Tech report ISI/RR-87-180, University of Southern California, 1987.
2. H. Takagi and L. Kleinrock, Optimal transmission ranges for randomly distributed packet radio terminals, *IEEE Transactions on Communications*, 32(3):246–257, March 1984.
3. R. Bellman, On a Routing Problem, *Quarterly of Applied Mathematics*, 1(16):87–90, 1958.
4. L.R. Ford Jr., Network flow theory, The RAND Cooperation, Santa Monica, California, Technical report P-923, August 1956.
5. C. Perkins, E. Belding-Royer, and S. Das, Ad hoc on-demand distance vector (AODV) routing, RFC 3561, IETF, July 2003.
6. T. Clausen and P. Jacquet, Optimized link state routing protocol (OLSR), RFC 3626, IETF, October 2003.
7. E. Kranakis, H. Singh, and J. Urrutia, Compass routing on geometric networks, in *Proceedings of the 11th Canadian Conference on Computational Geometry (CCCG*

'99), Vancouver, Canada, August 1999, pp. 51–54.

8. I. Stojmenovic and X. Lin, Loop-free hybrid single-path/flooding routing algorithms with guaranteed delivery for wireless networks, *IEEE Transactions on Parallel and Distributed Systems*, 12(10):1023–1032, 2001.
9. R. Nelson and L. Kleinrock, The spatial capacity of a slotted ALOHA multihop packet radio network with capture, *IEEE Transactions on Communications [legacy, pre-1988]*, 32(6):684–694, 1984.
10. P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, Routing with guaranteed delivery in ad hoc wireless networks, *Wireless Networks*, 7(6):609–616, 2001.
11. G.T. Toussaint, The relative neighborhood graph of a finite planar set, *Pattern Recognition*, 12:261–268, 1980.
12. K. Gabriel and R. Sokal, A new statistical approach to geographic variation analysis, *Systematic Zoology*, 18:259–278, 1969.
13. J.A. Bondy and U.S.R. Murty, *Graph Theory with Applications*, Macmillan, London, U.K., Elsevier, North-Holland, 1976.
14. H. Frey and I. Stojmenovic, On delivery guarantees of face and combined greedy-face routing algorithms in ad hoc and sensor networks, in *Proceedings of the 12th annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '06)*, Los Angeles, CA, September 2006, pp. 390–401.
15. J. Gao, L.J. Guibas, J. Hershberger, L. Zhang, and A. Zhu, Geometric spanner for routing in mobile networks, in *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '01)*, Long Beach, CA, 2001, pp. 45–55.
16. X.Y. Li, G. Calinescu, and P.J. Wan, Distributed construction of a planar spanner and routing for ad hoc wireless networks, in *Proceedings of the 21th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '02)*, 2002, pp. 1268–1277.
17. X.-Y. Li, I. Stojmenovic, and Y. Wang, Partial delaunay triangulation and degree limited localized bluetooth scatternet formation, *IEEE Transactions on Parallel Distributed Systems*, 15(4):350–361, 2004.
18. P. Boone, E. Chavez, L. Gleitzky, E. Kranakis, J. Opatrny, G. Salazar, and J. Urrutia, Morelia test: Improving the efficiency of the gabriel test and face routing in ad-hoc networks, *Lecture Notes in Computer Science*, 3104:23–34, 2004.
19. S. Datta, I. Stojmenovic, and J. Wu, Internal node and shortcut based routing with guaranteed delivery in wireless networks, *Cluster Computing*, 5(2):169–178, April 2002.
20. J. Zhao and R. Govindan, Understanding packet delivery performance in dense wireless sensor networks, in *Proceedings of the First International Conference on Embedded Networked Sensor Systems (SenSys '03)*, Los Angeles, CA, 2003, pp. 1–13.
21. A. Woo, T. Tong, and D. Culler, Taming the underlying challenges of reliable multihop routing in sensor networks, in *Proceedings of the First International Conference on Embedded Networked Sensor Systems (SenSys '03)*, 2003, pp. 14–27.
22. J.A. Sánchez and P.M. Ruiz, Locally optimal source routing for energy-efficient geographic routing, *Wireless Network (WINET) Journal*, November 2007.
23. M. Mauve, H. Füßler, J. Widmer, and T. Lang, Position-based multicast rout-

ing for mobile ad-hoc networks, Department of Computer Science, University of Mannheim, Technical report TR-03-004, March, 2003.

24. M. Transier, H. Füßler, J. Widmer, M. Mauve, and W. Effelsberg, Scalable position-based multicast for mobile ad-hoc networks, *First International Workshop on Broadband Wireless Multimedia: Algorithms, Architectures and Applications (BroadWim 2004)*, San Jose, CA, 2004.
25. J.A. Sánchez, P.M. Ruiz, J. Liu, and I. Stojmenovic, Bandwidth-efficient geographic multicast routing protocol for wireless sensor networks, *IEEE Sensors Journal*, 7:627–636, September 2007.
26. M. Heissenbüttel, T. Braun, T. Bernoulli, and M. Wächli. BLR: Beacon-less routing algorithm for mobile ad-hoc networks *Elsevier Journal of Computer Communications*, 27(11):1076–1086, July 2004.
27. B. Blum, T. He, S. Son, and J. Stankovic, IGF: A state-free robust communication protocol for wireless sensor networks, Department of Computer Science, University of Virginia, Technical Report, 2003.
28. H. Füßler, J. Widmer, M. Käsemann, M. Mauve, and H. Hartenstein. Contention-based forwarding for mobile ad hoc networks, *Ad Hoc Networks*, 1(4):351–369, 2003.
29. M. Chawla, N. Goel, K. Kalaichelvan, A. Nayak, and I. Stojmenovic, Beacon less position based routing with guaranteed delivery for wireless ad-hoc and sensor networks, in *Proceedings of 19th IFIP World Computer Congress (WCC '06)*, Santiago-de Chile, Chile, August 2006.
30. M. Witt and V. Turau. BGR: Blind geographic routing for sensor networks, in *Proceedings of 3rd Workshop on Intelligent Solutions in Embedded Systems (WISES '05)*, Hamburg, Germany, May 2005, pp. 51–61.
31. J.A. Sánchez, R. Marin-Perez, and P.M. Ruiz, BOSS: Beacon-less on demand strategy for geographic routing in wireless sensor networks, in *Proceedings of 4th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS '07)*, Pisa, Italy, October 2007.
32. R.M. Karp, Reducibility among combinatorial problems, *Complexity of Computer Computations*, 43:85–103, 1972.
33. P.M. Ruiz and A.F. Gomez-Skarmeta, Approximating optimal multicast trees in wireless multihop networks, in *Proceedings of 10th IEEE Symp. on Computers and Comms. (ISCC'05)*, La Manga del Mar Menor, Spain, June 2005, pp. 686–691,
34. E.M. Royer and C.E. Perkins, Multicast operation of the ad-hoc on-demand distance vector routing protocol, in *Proceedings of 5th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, Seattle, WA, August 1999, pp. 207–218.
35. S. Ju Lee, W. Su, and M. Gerla, On-demand multicast routing protocol in multihop wireless mobile networks, *Mobile Networks and Applications*, 7(6):441–453, 2002.
36. P.M. Ruiz, A. Gomez-Skarmeta, and I. Groves, The MMARP protocol for efficient support of standard IP multicast communications in mobile ad hoc access networks, in *Proceedings of International Mobile IP-based Network Developments (MIND) Workshop*, London, U.K., October 2002.
37. J.G. Jetcheva and D.B. Johnson, A performance comparison of on-demand multicast routing protocols for ad hoc networks, School of Computer Science, Carnegie

Mellon University, 2004.

38. M. Witt and V. Turau, The impact of location errors on geographic routing in sensor networks, in *Proceedings of the 2nd International Conference on Wireless and Mobile Communications (ICWMC'06)*, Bucharest, Romania, July 2006.
39. S. Giordano, I. Stojmenovic, and L. Blazevie, Position based routing algorithms for ad hoc networks: A taxonomy, *Ad Hoc Wireless Networking*, 103–136, 2004.
40. J. Li, J. Jannotti, D.S.J. De Couto, D.R. Karger, and R. Morris, A scalable location service for geographic ad hoc routing, in *Proceedings of the 6th annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '00)*, Boston, MA, 2000, pp. 120–130.
41. J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, Energy-efficient broadcast and multicast trees in wireless networks, *Mobile Networks and Applications*, 7:481–492, 2002.
42. J.A. Sánchez and P.M. Ruiz, LEMA: Localized energy-efficient multicast algorithm based on geographic routing, in *Proceedings of the 31st IEEE Conference on Local Computer Networks (LCN '06)*, Tampa, FL, November 2006, pp. 3–12.
43. H. Frey, F. Ingelrest, and D. Simplot-Ryl, Localized minimum spanning tree based multicast routing with energy-efficient guaranteed delivery in ad hoc and sensor networks, Institut National de Recherche en Informatique et en Automatique (INRIA), France, Technique Report RT-0337, June 2007.
44. H. Frey and F. Ingelrest, MFACE: A multicast backbone-assisted face traversal algorithm for arbitrary planar ad hoc and sensor network topologies, in *Proceedings of International Workshop on Theoretical and Algorithmic Aspects of Sensor and Ad-hoc Networks*, Miami, FL, 29 June 2007.

第9章 无线传感器网络中的媒体访问控制协议

媒体访问控制协议（MAC）定义了访问和控制共享媒介的规则，并且在有效和公平分享无线带宽方面起到重要的作用。无线频道的特性带来了新的问题，诸如位置依赖载波侦听、时变信道和突发错误，低功耗的要求提出了新的挑战。无线MAC得到研究人员的广泛关注，已经提出来了一些协议。这些协议是基于不同类型的结构、不同的应用和不同的媒介而设计的。

无线传感器网络（WSN）通常是由电池提供能量的，为了使得它的应用更加经济可行，这种网络需要在不充电或者不更换电池的情况下运行很长一段时间（例如数年）。因此，开发一项尽可能地延长电池的使用寿命的技术是非常重要的。因此，在WSN中能源效率可能会成为最重要的问题。

在WSN中对能量有效操作的需求促使在通信栈的所有层开发新的协议。假如在一个典型的传感器节点中无线接收器是最耗能的组件，在MAC协议控制无线收发器单元使用的链路层能够获得最大的收益。

传感器网络的MAC协议和传统的无线网络MAC协议在很多方面有很大的不同。WSN的MAC协议必须含有内置电源保护、移动性管理和故障恢复策略。而且，传感器MAC协议应该取得延迟和吞吐量这两种执行性能的均衡，最终降低能量消耗，从而最大化网络寿命。即总体上，通过无线电工作周期来获得。

本章介绍无线MAC协议的基本原理和概念，解释具体要求和设计一个特定的WSN MAC协议的约束条件。此外，还介绍了归类和讨论一系列典型的专门为WSN而设计的MAC协议。最后总结，我们提出研究方向和确定将来媒体访问研究的一些悬而未决的问题。

9.1 简介

微机电系统的研究进展，低功耗高集成的数字电子产品，微小的微处理器和低功率无线电技术创造了低成本、低功耗和多功能的传感器设备，它们能够观察并对它们周围环境的变化作出反应。传感器设备装备有一个小的电池、一个无线接收器和一套用于收集周期环境信息的传感器。这些传感器的出现使得工程师去设想网络是由大量部署在人们感兴趣的宽阔的区域的[1-6]的传感器组成。一个典型的WSN由大量的传感器设备合作完成一项共同的任务，例如环境监测和使用无线电向中心节点（sink节点）报告所收集的数据。WSN能够应用在很多的民用和军用领域，

包括在战场的目标跟踪[7]、环境监测[8,9]、民用建筑监测[10]和工厂维修[11]。在许多应用中传感器节点在没有精心的准备之前以一个特定自组织方式部署[12]。它们必须自己组织形成一个多跳，无线通信网络从而能够彼此通信，并且有一个或者多个 sink 节点[13]。因为大量的节点部署密集，邻居节点之间可能非常接近。因此，在传感器网络中希望多跳通信比传统的单跳通信消耗更低的能量。另外，发射的功率水平可以保持在较低的水平，这在隐藏操作中是非常重要的。多跳通信也能够有效地克服一些在长距离无线通信下的信号传播的影响。尽管多跳通信模式是其中的优点之一，但是正如参考文献［107］所讨论的，多跳方案并不总是最好的解决方案。为了控制传感器网络的操作，通过一个控制中心（sink），远程用户可以发布命令到传感器网络中，sink 负责分配数据收集、数据处理和转移任务到传感器，以及此后可以通过 sink 接收到感知数据。

假设传感器节点装备有有限的、通常不可替代的电源能源，WSN 必须具有一个内部的均衡机制使得传感器网络能够省电，并且使得终端用户在更低的吞吐量和更高延时的代价下延长网络的生命周期[1]。

传感器节点的能量限制和对于 WSN 上能量的有效操作的需求激发了传感器网络上的大量研究，从而导致了在开放系统互联（OSI）的通信栈的所有层开发一些新的通信协议。鉴于在一个典型的传感器节点中无线电是最能消耗能量的部分，在 MAC 协议控制无线电单元使用的链路层能够获得巨大的收益。

在传统的无线网络中 MAC 协议已经被广泛地研究。时分多址（TDMA）、频分多址（FDMA）和码分多址（CDMA）是被广泛地应用于现代移动通信系统的 MAC 协议。它们的基本思想是通过分别调度节点的时间、频率和正交码到不同的子信道，从而来避免干扰。因为这些子信道之间不会相互干扰，在这个组中 MAC 协议通常是没有碰撞的。

另外一类 MAC 协议是基于竞争的。不同于预先分配通信，节点共享相同的信道。在这样的系统中碰撞发生在竞争过程中。典型的基于竞争的协议的例子是 ALOHA[14]和 CSMA[15]。在 ALOHA 中，当产生新的消息包时（纯 ALOHA）或者在下一个可用的时隙（时隙 ALOHA）[16]时，节点简单地传输一个包。碰撞的消息包被丢弃或者重发。在 CSMA 中，节点在发送数据之前侦听信道。假如它检测到一条繁忙的信道，会延迟访问并且过段时间重试。

WSN 很多特征与传统的无线网络有很大的不同。这些特征使得传统的 MAC 协议不适合 WSN，激发了在设计 MAC 协议领域的大量研究。

在参考文献[1,3,17-31,80,81]中讨论了 WSN 中 MAC 协议的很多方面。然而，考虑到无线传感器网络自身的特点，本章全面回顾传感器网络的 MAC 协议需要克服的最近发展的和具有挑战性的问题，并讨论文献中一系列典型的解决方案。此外，我们对于现在还没有研究的和没有深入研究的问题给出未来的研究方向。

9.2 无线传感器网络

WSN是一个为一些潜在的应用，如环境监测、监视、军事、健康、安全等问题而设计的独立存在的自组织网络。一个典型的WSN是由大量的节点组成的，密集部署在环境区域内部或者非常接近于环境区域的地方。一个传感器节点由下面四个基本的部分组成，如图9-1所示，一个感知单元，一个处理单元，一个接收单元和一个能量单元；也有可能根据应用的需要增加一些模块，例如位置管理单元[1]，一个流动性管理单元和一个发电单元。一旦节点部署到目标区域，它们自动地从环境中收集数据以及建立一个自组织网络发射所收集的数据到基站（sink节点）。基站聚集和分析所收集到的数据，并且判断在部署区域是否有异常情况发生。

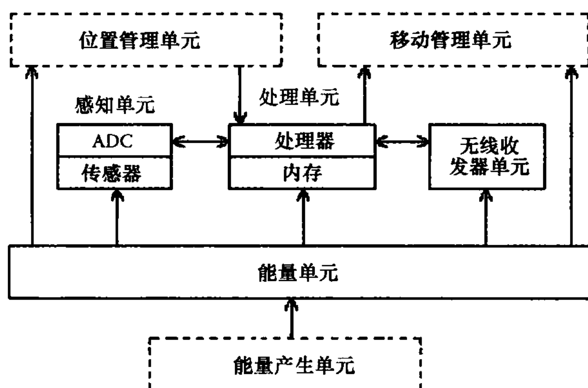


图9-1 一个传感器节点的组成

通常在传感器网路中的传感器节点是由电池供电的，应该有几个月或者几年的寿命。对网络部署后的成千上万个传感器电池的更换或者更新电源变得非常的困难或过于昂贵。因此，在WSN中能源效率成为WSN的一个关键的设计挑战。本节介绍和讨论了使得WSN不同于传统无线网络的具体特性。

9.2.1 无线传感器网络特性

WSN在很多方面和移动Ad Hoc网络极其相似，但是无线传感器网络有一些鲜明的特征不同于其他的网络，并且由此带来了一些具有挑战性的问题。这里我们仅仅讨论三个特别感兴趣，并且影响传感器网络协议设计的特性。

有限的资源。传感器节点具有有限的能量、内存和计算能力。因此，任何为WSN设计的算法不能假设能量是不受限的，而且必须节约地使用仅存的有限的资源。

传感器节点容易失效。由于内部的不稳定性和传感器的能量限制，传感器节点容易失效。因此，我们非常有必要知道在网络内部哪些节点或者哪一个区域的损耗是最高的。这些信息对于设计容错协议或者监测机制具有潜在价值，因此问题区域可能被重新部署，关键数据被重新路由来避免这些区域数据的高的丢失率。每个节点的一些可能应用的信息损耗会导致简化数据流或者增加大规模传感器网络的不稳定性。

无线带宽的限制。因此，在设计传感器网络协议中不能依赖于使用现行的、没有扩展的或者带宽有效的知识。这使得在传感器网络中直接收集损失率数据是不可能的。而且，由于带宽的限制，单个传感器节点收集和传输损失率数据到一个集中的位置进行处理是不可行的。

9.2.2 传感器节点的功耗

如上所述，对于一个传感器节点的能量供应是非常珍贵的：电池的容量很小，通过能量净化再充电的方法是复杂多变的。因此，必须严格控制传感器节点的能量消耗。能量消耗的主要部分是感知单元、数字处理单元和无线电收发单元。

有助于减少这些组件能量消耗的一个重要的方法来自于芯片层和能源技术。设计一个低能量级的芯片对一个节能的传感器节点是最好的开端。但是这仅仅是成功的一半，当模块操作不当的时候，这些设计者所取得的优势会被轻易地抵消。对于正确操作的一个重要的发现是在大部分时间内传感器节点没有什么可做。因此，最好是将其关闭。自然地，当发生外部刺激事件或者时间周期到达时，它应该重新被唤醒。因此，完全地关闭一个节点是不可能的，而是将它的运行状态切换到低功耗状态。引进和使用减少耗能的工作状态来换得降低的功能性是节能无线传感器节点的核心技术[82]。下面，我们将更加深入的了解传感器节点主要的能量消耗者的能量消耗。

（1）感知单元　感知单元由环境传感器和模-数（A-D）转换器组成，把物理现象转换为电信号。在感知单元中有几个能量消耗的来源：信号采集和把物理信号到电信号的转变，信号调整和模-数转换器。这个单元的能量消耗是相对的固定的，并且依靠增加集成和高效的模拟电路设计改善其能源效率。已经被验证的是被动传感器，例如温度，地震等，相对于传感器节点的其他模块其能量消耗可以忽略不计[86]。

（2）数据处理单元　在传感器节点中采用的大部分数字电路通常用于命令和控制功能，基带信号处理单元和该协议的执行堆栈。在数字电路中的能量消耗是由静态和动态的功耗所决定的。动态的功耗是指转换电池容量的结果以及门控电源电压。相对于动态功耗，静态功耗的主要能量消耗在数字电路单元。减少静态功耗的最简单的方法是关闭空闲状态的能源供能，这就是所谓的电源门控。但是应该注意的是关闭复杂电路可能导致时间和能量的消耗。解决这些问题以及节约能源的方法

超出了本章的范围。详细的信息见参考文献［83］。

（3）无线电收发单元　在系统操作中无线通信是最主要的能量消耗者。很难推断出通信系统的能量消耗，因为有很多的变量影响着系统的执行性能。总体上，无线电收发单元的能量消耗有两部分组成：1）取决于传输的距离和调制参数的RF组件。2）一个占有能量的电子元器件，而能量则由于电路执行频率合成、滤波和升频转换等原因被消耗。一个完整的无线电传输所消耗的平均能量能够由下面的方程描述[38,85]：

$$E = P_{tx}(T_{transmit} + T_{start\text{-}up}) + P_{out}T_{transmit} \tag{9-1}$$

式中　P_{tx}——发射机所消耗的功率；

$T_{transmit}$——实际的传输时间；

$T_{start\text{-}up}$——收发器所启动的时间；

P_{out}——驱动天线所输出的发射功率。

由于在传感器网络中的低数据率，在空闲周期中无线电所传送的数据包可能会很小。启动时的功率将占据主动传输时能量的主要部分。当无线的高偏置电流需要节点接受关闭代价，节点应该分摊启动时的能量到更多的传输比特中以减少每一个传输比特的能量代价。

除了上面提到的因素，在无线电还有很多其他的因素导致能量损耗。总体上，无线电在下面四个不同的运行模式下进行操作：发射、接收、空闲以及休眠。一个重要的影响因素是随着无线电操作模式的改变消耗了大量的能量。表9-1显示了在休眠模式下大多数的无线操作系统导致了显著的高能耗[58]，几乎相当于在接收模式下的能量消耗。显然，当没有接收或者发送数据时完全关闭无线电而不是把其转换到休眠模式是非常重要的。然而，由于启动时的能量消耗，非常频繁地打开和关闭一个开关有时可能会比让收发单元进入休眠模式消耗更多的能量。而且，随着传输的数据包的尺寸变得更小，过渡能量而不是在接收和传输数据之间所消耗的能量[38,85]会占主要部分。因此，当设计一个高效能耗的MAC协议时考虑到这个问题是非常重要的。

表9-1　一个标准的无线电能量消耗

无线电模式	能量消耗/MW	无线电模式	能量消耗/MW
发送	14.88	空闲	12.36
接收	12.50	睡眠的	0.016

9.2.3 通信模式

一般来说，大多数的WSN应用含有自身区别于传统的无线网络的显著特征。一个普遍的特征是传感器节点仅仅部署去监测环境以及报告数据到管理中心以进行

进一步的数据处理。监测程序不需要进行大量的传输或者数据处理信息。在 WSN 中负责产生通信的通信模式是极不对称的，在 sink 节点（汇聚节点）附近出现显著的更多的流量。然而，在 WSN 中可以定义三个通信模式：广播、收敛计算和局部散播[87]。广播通信模式主要被应用在基站（sink）上去传输一些信息到网络中的所有传感器节点。广播的信息可能包括有关传感器查询处理架构的查询，传感器节点的程序更新，以及控制整个系统的控制包。广播类型的通信模式不应该和广播类型的数据包混淆。对于前者，将接收网络中的所有节点的包，而后者将接收节点通信范围内的节点的包。

在某些情况下，检测到干扰源的传感器会在局部之间进行互相的通信。这种通信方式叫做局部散播，这里传感器传递信息到其在一定的范围内的邻居节点。然后，检测到事件的传感器需要传送它的感知信息到信息中心。这种通信模式叫做收敛计算，这里一组传感器将数据传输到一个特定的传感器。目标节点可以是一个簇头、数据融合中心或者基站。

在基于分聚的协议中，簇头和它们的成员之间进行通信，因此预定的接收器可以不是簇头的所有邻居，而是一个邻居的子集。为了满足这种情况，定义了第四种通信模式——多播通信模式，这里传感器传送信息到特定的传感器子集。

9.3　无线 MAC 协议的概念和基本原理

在这节中，我们简要的讨论了无线 MAC 协议一些基本的概念和方面，因为在 WSN 中使用的协议含有很多的问题，并且对于这个更加广阔的领域已经有很多现存的方法。

在最近的几十年里无线 MAC 协议已经获得了研究者和商业开发者很大的重视，且已经有了很多的文献[88]。

在开发 WSN 以前，对于早期的 MAC 协议研究者来说能源方面不是最需要考虑的。但是今天，尤其是在 WSN 中能量对于所有的设计的元素来说是至关重要的。

9.3.1　无线 MAC 协议的需求和设计条件

对于传统 MAC 协议最重要的性能要求是吞吐率，高信道利用率，可靠性，稳定性，公平和低延迟性，以及低开销。MAC 协议的开销来自于每一个数据包的代价，碰撞或者额外的控制数据包交换。假如 MAC 协议允许两个或者多个节点同时发送数据可能会发生碰撞，这样有可能导致接收者无法正确地解码数据包，导致上层做一个包重传。在实际的应用当中，对于传递时间或者最小可用数据率提供确定性的或者随机的保证是非常重要的。有时，利用优先级概念使用一些级别的服务区分比为更重要的包优先提供服务更能够满足用户需求。MAC 协议的操纵和执行性

能在很大的程度上受到了底层物理层的影响。因为WSN使用无线媒介，它含有无线通信所有知名的问题。普遍的问题是时变比特率和高误码率，这是由快速和慢速的衰落，路径损耗，衰减或者热噪声等物理现象所引起的。此外，使用的调制方案、频率、接收和发送者之间的距离，以及传播环境等对于比特误码率有很大的影响。最后，设计MAC协议在很大的程度上依靠于预期的传输负载模式[87]。

9.3.2 无线MAC协议的分类

无线MAC协议总体上可以被分成三个不同的类型：固定分配协议、随机分配协议和按需分配协议[88]。

固定分配协议，把信道带宽分为固定的静态方式，独立的信道活动。FDMA、TDMA、CDMA和空分多址（SDMA）是这个类的常见形式。

随机分配协议，整个带宽作为一个单一的能够被随机访问的实体提供给用户。ALOHA和CSMA属于这一类。

按需分配协议，在用户中需要交换明确的控制信息。这些协议可以进一步的划分为中央控制，例如投票或者探询协议和分布式控制，例如多业务接入平台（MSPA）。

9.4 无线传感器网络的介质访问

在这节当中，我们缩小对WSN中MAC协议所讨论的具体要求和设计考虑。在WSN中引起能量消耗的源也会在这节中被研究。

9.4.1 在无线传感器网络中的能源资源消耗

在传感器网络中延长传感器节点的寿命和尽可能长时间地保持网络操作的可用是至关重要的。节能的MAC协议应该考虑多种造成的资源浪费和使得传感器电池自身消耗得非常快[33-36]的原因。在这节中我们罗列和讨论一些在设计MAC协议时应该考虑的一些问题。

数据碰撞是能量浪费最主要的原因。当两个数据同时被传输并且相互碰撞，它们已经被损坏且必须被丢弃，这些数据包的重传需要增加能量的消耗。另一个在无线中引起能量浪费的重要原因是串扰。串扰意味着一个节点接收到目的地为其他的节点的数据包。尤其是在流量负荷比较重的环境中和密集的网络串扰是引起能量浪费的一种重要的因素。因为很多物理参量的传感范围比通信范围要小得多，所以密集的传感器网络部署非常常见。

一个典型的无线单元能够在四种不同的模式下运作；空闲、接收、发送以及睡眠。正如所期望的，无线消耗比较多的能量在发送和接收模式，运行在空闲模式同样会消耗能量，尤其是在感应器是空闲的时候没有数据发送，这种情况通常叫做侦听空闲。完全关闭无线电而不是去转换模式到空闲状态是非常可取的。然而，频繁

的进行模式之间的转换，尤其是把睡眠状态转换为激活状态，会因为启动时需要消耗能量而导致无线接收单元将会比处于休眠模式下浪费更多的能量。

控制包代价同样是我们这里考虑的能量消耗的主要部分。发送、接收和侦听控制数据都消耗能量。由于控制数据包不会直接传送有效的应用数据；它们同样减少了有效的流量。用于建立数据传输的控制数据包应该尽可能低。在传感器网络中避免多余的发送会提高能量效率。多余的发送是当目的地节点是休眠状态或者没有准备接收时传输信息引起的。这样同样导致了系统能量资源的浪费，因此应该被避免。

传输波动，在 WSN 的一些应用产生的传输在空间和时间上波动，结果在高峰负荷处可能驱动传感器网络进入拥塞导致碰撞的可能性提高，因此，在随机后退程序过程中很多的时间和能量是浪费在等待中的[23]。

从能量的观点来看选择适当大小的包尺寸也是一个重要的问题。随着包尺寸变得越来越小，相比于在接收和发送数据包之间所消耗的能量[38,85]转变的能量变得更加显著。

大多数这些额外的开销是由基于竞争技术的 MAC 协议所引起的。当转换到基于调度技术的 MAC 协议时，例如 TDMA，可能首次会非常的吸引人，因为传感器节点是预定的，以及在没有任何的数据传输之前每一个节点清楚地知道那一个点应该发送和接收、空闲监听，额外开销和碰撞根本就不会发生。但是，这些以复杂协议为代价的优势会降低处理传输波动的灵活性以及网络拓扑结构的改变显著地增加了协议开销。解决这些问题的方案之一是去应用某种超储备和使用大规模的框架去处理峰值负荷。另外一种方法是动态调整框架尺寸，但是这会大大增加协议的复杂性，因此，不适用于资源有限型设备，例如传感器节点。

WSN 硬件和通信协议的设计是通过避免或者减少能量浪费去实现最低的能量消耗这个目标的。一个完整的能源管理方案必须不仅要考虑无线电，而且还要考虑传感器节点的所有消耗能量的硬件单元。

然而，在 MAC 层面上，能源的效率能够通过避免或者减少空闲侦听，重发不必要的监听和过多发送来提高。当不需要的时候关闭无线对于节约能源来说是非常重要的举措。

9.4.2　无线传感器 MAC 设计需求和权衡

传感器节点是由电池供电且因为空间和成本的限制使用大电池来供电是不可能的。再者，定期更换电池也是不可行的。因此必须使得传感器节点能够尽可能地节省更多的能量，延长网络的寿命。由于在传感器节点中无线电设备通常是最消耗能量的，通过控制无线电操作可以节省大量的能量。一个能量高效的 MAC 协议因为其直接控制无线单元操作能够很大地减少无线电单元的能量消耗。

MAC 协议受到很多因素的影响。一个设计良好的 MAC 协议应该考虑设计一系

列的执行特性和在它们之间的权衡。无线传感器 MAC 协议所需要的最重要的执行特性是[17,32]：

（1）避免碰撞　它是所有 MAC 协议的首要任务。它决定了一个节点何时以及如何访问媒介以及发送数据。在正常操作下碰撞并不是被完全避免的；基于竞争的 MAC 协议接受一定水平的碰撞。但是所有的 MAC 协议应该避免频繁的碰撞。

（2）能源效率　如前所述，能量对于传感器网络能源是一种稀缺的资源，而无线电作为传感器节点电源的最大消耗者，特别是在远程的传输和无线电一直保持开启时。因此能量感知 MAC 协议能够通过限制潜在的碰撞发送和接收能量，尽量减少使用控制消息，利用大部分可用的频率带宽去减少传输和接收能量，在休眠状态下把无线电调节到最低能量睡眠状态，最终避免在激活和睡眠状态之间过多的转换。

（3）可扩展性和适应性　它是与 MAC 协议密切相关的属性，能够适应网络规模、节点密度以及拓扑的改变。这些问题背后的一些原因是有限的节点寿命，增加新的节点到网络中和不同程度的干扰可能改变连接并最终改变网络拓扑结构。一个好的 MAC 协议应该能处理和适应这样的网络改变。

（4）延迟　它是发送者发送一个包到包成功地被接收者接收所需要的时间。在传感器网络中，延迟的重要性取决于应用。在某些特定的 WSN 应用中，检测到对象必须迅速地做出反应[84]。

（5）可靠性　可靠的数据传输是所有网络结构的基本设计目标。可靠性问题在 WSN 中具有重大意义，必须保证 WSN 应用中的数据包被成功地传递。在无线网络中，数据包丢失主要是由于缓冲区溢出和信号干扰。缓冲区溢出能够通过在 MAC 协议中使用一个缓冲区管理决策，从而阻止超过最大缓冲区大小的数据包数来避免。通过传输优先级或者滤波和聚合能够实现对缓冲区的控制。我们知道相邻传感器节点的数据读数是非常相关的，采用数据过滤和聚合机制可能减少数据传输从而降低通信量，避免缓冲区溢出和保存能量。由于通过使用足够高的传输能量和在节点中媒体访问的预防竞争机制，信号干扰造成的丢包能够被降到最低。

在 WSN 中的移动性对 MAC 协议的设计构成了一个挑战。MAC 协议应该适应移动模式的变化，使其既能够适应高移动性的感知环境也能够适应低移动性的感知环境。

信道利用率是指整个信道的带宽在通信中如何被利用。在传感器网络中信道的使用通常是次要的目标。

吞吐量是指在给定的时间内从一个发送者成功地传输到接收者的数据量。在传感器网络中很多的因素影响吞吐量，包括碰撞避免有效性、延迟、信道利用率和控制开销。与延迟相同，吞吐的重要性也是和应用相关的。

公平性反映了不同的用户、节点或者应用公平地分享信道的能力。它是传统的语音和数据网络的重要属性。然而，在传感器网络中，所有节点为了一个单一的共

同任务合作。在某一个特定的时间，一个节点可能需要比另外的节点发送更多的数据。因此，不是平等地对待每一个节点，成功发送是对一个应用整体的执行性能的度量，而每个节点或者每个用户的公平性变得没有那么重要。

简而言之，上述属性反映了 MAC 协议的特点。对于 WSN，最重要的因素是有效的碰撞避免、能量效率、移动性、可扩展性以及对密度和节点的数量的适应性。其他属性通常是次要的。

9.5 无线传感器网络 MAC 协议的分类

根据用于访问共享信道的保守的机制，WSN 的 MAC 协议能够被分成三个大体的组：预定的、非预定的（或者是随机的）和混合协议。预定的 MAC 协议试着以一个有序的方式在传感器节点之间组织通信。最普遍的预定方法使用 TDMA 组织传感器节点，每个单一的传感器节点利用一个时隙。组织传感器节点通过同步和状态分发来提供降低碰撞和信息重传的能力。非预定的协议试图通过允许传感器节点在最低复杂度下独立操作去保存能量。尽管碰撞和空闲监听可能发生且会引起能量的损耗，非预定的 MAC 协议基本上不共享信息或者维持状态。混合的 MAC 协议综合了预定和非预定 MAC 协议的优势去补偿自身的弱点来提高协议的效率。混合 MAC 协议最大的优点是能够简单、快速地适应传输状态，这能够节省大量的能量，但是这种优势的获得是以协议复杂度为代价，限制了它的应用范围。一些已经提出的 MAC 协议不能够很容易地适应这种分类模式以及许多其他的现存的分类模式。然而，基于 MAC 协议设计之后的思想，下面会提供对每一类 MAC 协议更加具体的分类。

在这一节当中，由于篇幅的限制我们不能在文章中阐述所有的 MAC 协议，只是讨论一些典型的协议。

9.5.1 非预定的 MAC 协议

在无线网络中一个普通的 MAC 范例为 CSMA[15]。由于它的简易性、灵活性和鲁棒性，而非常受欢迎。它不需要很多的基础设施的支持：不需要无时钟同步和全局拓扑信息，动态节点加入和离去能够很好地控制而不需要额外的操作。然而，这些优势是以反复的试验为代价而摸索出来的——一个试验可能产生访问冲突，即超过两个“竞争”节点在同一时间发送数据，引起了目的地接收到的信号保真度下降。碰撞可能发生在一个节点的任何两跳邻居之中。尽管在传输数据之前进行载波监听能够很大程度地降低一跳邻居之间的碰撞，超越单跳载波监听机制不起作用。这导致了著名的隐藏终端/暴露终端问题，使吞吐量严重地下降，尤其是在高数据率传感器网络应用中。

参照图 9-2，隐藏终端问题可以按照如下解释：假设 A、B、C 和 D 是传感器节

点，其传输范围用一个圆来表示。节点 A 发起到节点 B 的传输。节点 C 未捕捉到节点 A 的传输，发起它到节点 D 的传输。两个传输在节点 B 碰撞。在暴露终端问题上，考虑到图 9-3 的例子；这里，节点 B 延迟其到节点 A 的传输，因为它侦听到传输到节点 C 到节点 D 的传输，尽管在节点 A 不会有碰撞。

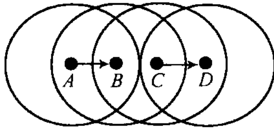


图 9-2　隐藏节点问题（环形代表传输范围和干扰范围）

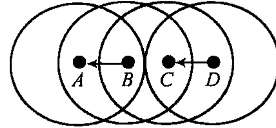


图 9-3　暴露节点问题（环形代表传输范围和干扰范围）

为了克服这些问题，需要使用碰撞避免技术，如使用 RTS/CTS（Request to Send/Clear to Send）握手。根据图 9-4，这个 RTS/CTS 握手工作原理如下所示：节点 A 发送 RTS，阻止在它的无线范围内所有节点的可能传输。节点 B 捕获到节点 A 的 RTS 后使用 CTS 应答。使用 CTS，节点 B 阻塞了它的邻居，让节点 A 发送数据。假如数据包被正确接收，节点 B 发送一个确认包到节点 A。

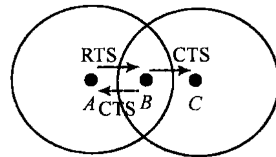


图 9-4　碰撞避免的 RTS/CTS 握手

非预定 MAC 协议有很多的优点。非预定 MAC 性协议按需分配资源；它们能够更加容易和灵活地适应节点密度和网络拓扑的改变，因为它们不需要获取当前的时间表或者加入另外一个传感器节点组。而且，非预定 MAC 协议同样允许传感器节点更容易地适应传输状态的改变，因为信道预定能够使用更细的微粒，并且传感器节点能够自适应地竞争信道。再者，非预定 MAC 协议不需要如 TDMA 协议一样的细粒度的时间同步技术。但是，由于传输是不协作的，非预定 MAC 协议有很多的弊端，如许多能量浪费、更高的碰撞率、空闲监听和串扰，另外，公平性在非预定性 MAC 协议中成为一个问题，因为不像在预定的 MAC 协议，非预定的 MAC 协议没有潜在的机制来平衡信道的使用。

基于 MAC 协议设计背后的思想，在这节中提供了对非预定协议更加详细的分类。

9.5.1.1　多通道的 MAC 协议

使用多个无线电接收器去保存一个传感器节点的能量在单个传感器节点中可能不是一个很好的选择，但是对于传感器节点，基于这种技术的一些设计方法能够显著地降低能耗。如果需要增加带宽或者响应时间，使用多个无线信道使得传感器节点能够在不同的信道上独立地同步通信。这个益处需要增加硬件。首先，无线收发器不断地消耗能量，即使是在睡眠状态，因此增加无线收发器增加了能量消耗，但

是却降低了节点整个能量消耗。其次，一个多无线收发器系统必须具有接收和处理来自多信道数据的计算能力。然后，多个无线收发系统相对于单一无线收发系统需要更高性能的通信机制和处理能力。PAMAS 协议是典型的这种类型的协议。

PAMAS：能量感知媒体访问协议和信号

PAMAS（The Power-Aware Medium Access Protocol and Signaling）协议是一个基于 CSMA 的协议，当节点不发射和接收数据时会自己断电[89]。这种方法需要节点对于控制和数据采用两个单独的无线频道。控制频道被用于信号交换，而数据信道用于定期的传输。使用两个信道最低限度地减少了潜在的碰撞。在 PAMAS 中，信息传输由源在控制信道上发送一个 RTS 信息到目的地发起。目的地检查数据和控制信道，然后决定是否发送一个 CTS。假如目的地在数据信道上没有探测到任何的活动，并且没有听到 RTS 或者 CTS 信息，那么它应答一个 CTS 信息。一个未及时接收到 CTS 信息的源将退而使用二进制指数算法。一旦源收到 CTS 信息，它通过数据信道传输数据信息。目的地一旦开始接收数据信息便在控制信道传输忙音，因此附近的节点意识到它们可能不能够使用数据信道。PAMAS 发送长度为 RTS 或者 CTS 信息两倍的忙音。而且，在数据接收阶段，目的地接收到一个 RTS 信息或者检测到控制信道上有噪声，就传输一个忙音来干扰 CTS 信息回复和阻止进一步的数据传输。图 9-5 表明了在 PAMAS 协议中的信息传输。

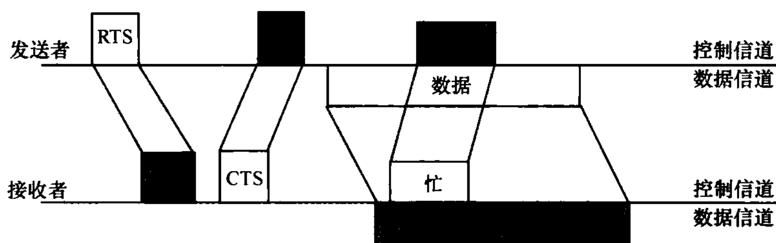


图 9-5 PAMAS 数据传输

发送者不能够建立连接时就转换为睡眠，稍后重试。节点停留在睡眠模式的时间是基于节点之间在控制信道上特殊探测信息之间的交换决定的。将未参与通信的节点转换为睡眠状态能够节省将近 70% 的能量。然而协议要求节点进行媒介监听来决定是否传输并不会完全消除碰撞。而且，协议要求节点对于独立的信道（控制信道和数据信道）有无线电单元，从而增加了设计传感器的成本、尺寸和复杂性。另外，控制访问两个无线媒介增加了 MAC 协议的复杂性。大多数传感器网络的特性就是数据信息太小，从而导致划分独立的数据和控制信道的优势降低。然而，那些通过 PAMAS 提出的思想可能对像多媒体传感器网络这种有大量数据信息的传感器网络起作用。

9.5.1.2 面向应用的MAC协议

在节约能量方面，应用特征可能被用于去提高MAC协议的有效性。例如，基于监测的传感器网络在大多数时间内具有很低的通信量，但是当需要关注的事情发生时可能产生相对很大的数据量。当传感器网络没有数据传输时，运行在基于假定的常量通信量上的MAC协议将节省能量。一个最近的和典型的利用应用特征去保存相当数量的能量的协议是CC-MAC。

CC-MAC：协作的MAC协议

通过利用位置邻近的传感器节点产生的测量数据是相关的这一事实，CC-MAC协议试着在满足应用的需求的同时去节省能量[41]。为了达到节约能源的目的，CC-MAC过滤来自高度相关的传感器节点的测量信息，从而尽量减少传感器网络需要处理的通信量。最小化通信量导致无线媒介间竞争的减少，因此减少了碰撞。

CC-MAC由两部分组成：事件MAC（E-MAC），过滤传感器节点的测量数据来降低通信量和网络MAC（N-MAC），将过滤掉的数据转发到sink节点。E-MAC通过仅仅允许那些至少被相关的距离所划分的传感器节点产生数据来降低某个区域的通信量。其他节点周期性地休眠以节约能量以及激活来转发信息。在整个网络中相关的传感器节点循环转换形成测量数据的角色来平衡能量消耗。N-MAC转发来自产生测量数据的传感器节点的数据到sink节点，但是因为E-MAC协议已经清除了多个测量数据中出现的大多数冗余数据，转发通信量变得更加的重要。

CC-MAC主要的缺点是要求传感器节点拥有或者获得关于邻居的范围信息，因此E-MAC能够从相关的传感器节点中过滤数据。此外，CC-MAC协议的复杂性会限制协议的应用。另外，随着感知事件数量的增加，特别是如果感知条件随着时间变化，整个网络中相关半径和分布的计算开销将增加。对于大型的网络，这个开销将变得非常显著。

9.5.1.3 多路径数据传输MAC协议

在这类MAC协议中，MAC协议利用后退机制来降低碰撞的概率。MAC协议经过一段延迟之后仅仅传输数据信息的多个副本，同时删除由控制信息和载波监听造成的开销。传输数据信息的多个副本增加了数据传输的可能性。鉴于此类协议的简单性，出现了很多的缺点和无效性。由于传输是未经协商的，很有可能发生碰撞。通过增加后退的时间间隔，这个问题能够解决，但是结果是增加了消息延迟。另外，尽管传感器节点没有交换任何有关数据发送成功的握手信息，协议在没有数据传递保证的情况下，通过多条路径传输相同信息的多个副本会浪费能量。然而，这种类型的协议对于产生轻量级的传输，并且仅仅需要发送有限信息到目的地的传感器网络来说是非常有用的。典型的例子如参考文献［90］。

SRBP、ARBP和RARBP

在参考文献［90］中作者提出了说三个协议，其中的每个都比前一个有所加强。简单随机后退协议（Simple Random Back-off Protocol，SRBP）是第一个协议，

其工作仅仅是在一个初始的随机后退之后传输信息。传感器节点既不感知信道也不交换任何控制信息。自适应随机后退协议（Adaptive Random Back-off Protocol, ARBP）根据传感器节点密度和当前的传输状态来调整最大的后退间隔而增加 SRBP 的性能。通过利用两个子协议获得传感器节点密度和传输状态：密度感知协议（$P_{density}$）和信息传输感知协议（$P_{traffic}$）。更多的资料见参考文献［90］。最后一个协议是自适应随机后退协议（Range Adaptive Random Back-off Protocol, RARBP），它使用与发送者和接收者之间距离相关的信息来调整随机后退时间间隔。远离发送者的节点有更高的概率去选择小的后退值，因此能够尽早的发送。这个方式减少了由于后退机制造成的延迟。

然而，这个协议继承了很多上面列出的缺点。

9.5.1.4　基于汇合的 MAC 协议

仅仅在两个节点同时供电的情况下它们之间的通信才有可能。因此需要一种方法允许节点之间的即时通信。这个方法通常叫做汇合模式。在无线节点之间有很多的方式去完成汇合。最受欢迎的方案是叫做循环接收机。在这个方案中，节点周期性地打开和关闭电源，使用一个信标方法来表达请求或者想要进行通信。这种类型协议的例子见参考文献［44］。

TICER 和 RICER 协议

初始传输循环接收（Transmitted Initiated Cgded Receiver, TICER）和初始接收循环接收（Receiver Initiated Cycled Receiver, RICER）是两个相似的协议[44]。TICER 协议使得拥有数据的传感器节点在一个感知周期内周期性地发送 RTS 控制数据包。接收者周期性地侦听无线信道，假如探测到一个 RTS 信息，它们回复一个 CTS 信息。从而传感器节点能够传输数据信息。RICER 反向操作，因此当它们从正常调度的休眠时间醒来时，接收者周期性地发送信标。一个有数据要发送的传感器节点保持在苏醒状态并且监听信道，直到从想要到达的目的地节点接收到一个唤醒信标。接收到唤醒信标以后，它开始发送数据包。在正常地接收了数据包之后，这个过程在结束时有一个从目的地节点发送到源节点的确认信号。这个协议的作者提到，一些协议参数，例如控制信息交换时间以及信道特征在协议的整个执行过程中起到了很重要的作用。图 9-6 和图 9-7 分别表示了 TICER 和 RICER 方案。

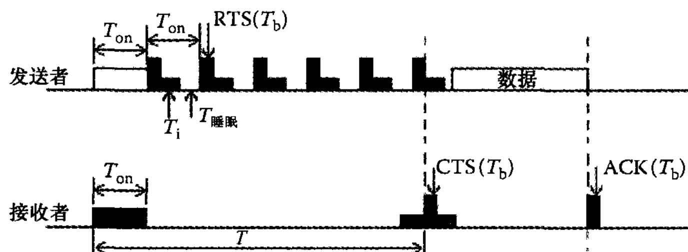


图 9-6　TICER 模式

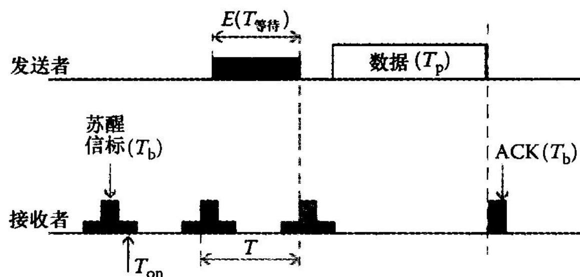


图 9-7　RICER 模式

9.5.1.5　基于前同步码的 MAC 协议

在非预定的 MAC 协议中，传感器节点可能不知道它们邻居的休眠时间表，因此它们必须以某种方法使用信息进行探测直到邻居激活。随着通信中的传感器节点及时地捕获彼此的信息，它们能够开始信息的发送。由前同步码技术节省的能量仅仅来自于当需要时，并且仅仅对于传输时期的邻近节点的同步中。然而，基于前同步码的能量节约在很大的程度上受到传输模式的影响。另外，由于延迟的增加，大多数基于前同步码的协议的长前同步码可能引起性能下降，从而限制了将这种协议应用到实时的或者延迟敏感的应用中。

伯克利 MAC

伯克利 MAC（B-MAC）协议中，基于目标的占空比，传感器节点在传感器网络中独立的遵循休眠时间表[43]。因为传感器节点独立地操作时间表，对于信息传输 B-MAC 使用非常长的信标或者前同步码。源传感器节点传输一个足够长的信标，目的地节点周期性地感应信道，并且有足够的时间醒来和感知活动。感知到信道活动的传感器节点仍然保持清醒来接收信标后的信息，或者如果它们未检测到信道上的活动就回到休眠状态。图 9-8 表示了在 B-MAC 中的信息传输。

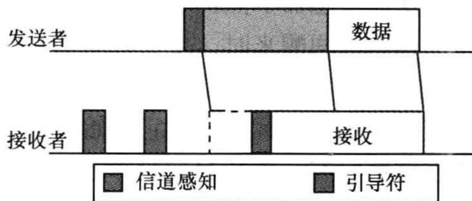


图 9-8　B-MAC 信息传输

B-MAC 是灵活的，这意味着通过协议接口，网络的设计者能够在协议中微调很多的操作变量，例如延迟和后退值。一方面，与传统的无线问题相比，B-MAC 没有提供任何的保护机制，例如隐藏终端问题。而且在 B-MAC 协议中的长前同步码可能引入额外的延迟，应该考虑到这个问题，可以通过传感器网络设计中的 B-MAC接口来控制。

9.5.2 预定的MAC协议

预定的MAC协议试图通过协调传感器节点使用一个共同的时间表来降低能量消耗。自TDMA出现以来，大多数提出的协议使用了一些其他形式的多址接入，例如频率（FDMA）或者码分（CDMA），增加了传感器节点的成本和能量需求。通过产生一个时间表，MAC协议规定了任何时间哪个传感器节点应该使用信道，从而限制甚至消除了碰撞、空闲监听以及串扰。未参与消息通信的节点将有可能进入一个休眠模式直到它们有工作要执行或者需要接收消息。另外，MAC协议能够共享传输量或者状态信息，因此单独的传感器节点能够在传感器节点的集合上最优化能量消耗，而不是仅仅对于一个单一的传感器节点。然而，这些优势是以增加消息来创建和维持一个时间表为代价的。节点移动，节点重新部署和节点死亡都使得调度维护更加复杂。进入网络的传感器节点必须等待直到它们了解到了时间表，以及可能进入时间表去使用信道。另外，当传感器节点死亡以及邻居传感器节点重新分配它的资源，都会导致延迟，因此一些源可能变得无用从而导致了没必要的延迟或者数据包丢失。那么，时间同步对于一个预定的协议是一个非常重要的问题，可以通过一个周期性的信标来产生，这会增加了收发器的使用，也可以使用高精度振荡器来实现同步，但这会增加传感器节点的成本。预定的MAC协议必须最小化额外的延迟和有限的吞吐量。典型地，增加的传感器节点仅仅在一个可能的时间段内的一小部分时间内能够访问无线信道。对于一个基于TDMA的MAC协议传感器节点能够访问信道的时间高度依赖于时隙的长度。典型地，只有一个传感器节点在间隔时间内能够传输，因此任何没有使用的时间被浪费了。减少时隙的长度可以降低浪费，但是在没有分片的情况下也会降低最大消息的长度。几个基于预定的MAC协议试图以共享消息中的附加信息或者以更高的工作周期为代价来克服传输量和延迟的限制。

9.5.2.1 基于竞争的分时隙MAC协议

分时隙竞争MAC协议通过让节点形成一个公共的睡眠/监听模式，允许它们以任意低的工作周期来使用无线收发器来节省能量。分时隙协议把时间分成为帧，每一帧被细分为一定数量的分时隙。每一帧的开始，有数据要发送的传感器节点醒来，开始竞争信道。因为所有通信都被分入时隙的监听部分，这个信道竞争导致的数据包碰撞的可能性很高。为了解决这个问题和提高分时隙MAC协议的性能，使用如RTS/CTS信号交换工具来避免碰撞。作为这种类型协议典型的例子，我们说明和讨论了S-MAC协议[32,33]。其他提到的协议是S-MAC协议的扩展，例如DMAC、TMAC、DSMAC、MS-MAC和ACMAC分别在参考文献［45-48，64］中提到。

感知MAC协议

感知MAC协议（Sensor-MAC，S-MAC）最基本的思想是基于周期性地休眠-侦

听计划和本地管理同步[32,33]。邻居节点形成虚拟簇去建立一个普遍的休眠计划。假如两个邻居节点存在于两个不同的虚拟簇中，它们在两个簇的侦听周期中被唤醒。S-MAC算法的一个缺点是可能遵循两种不同的时间表，因此它的空闲侦听和串扰会导致更高的能量消耗。通过周期性的SYNC数据包传播到最近的邻居来完成时间表的交换。通过载波监听来避免碰撞。而且，RTS/CTS数据包交换被用于单播形式的数据包。图9-9表示了一个发送者（A）/接收者（B）之间的通信的例子。

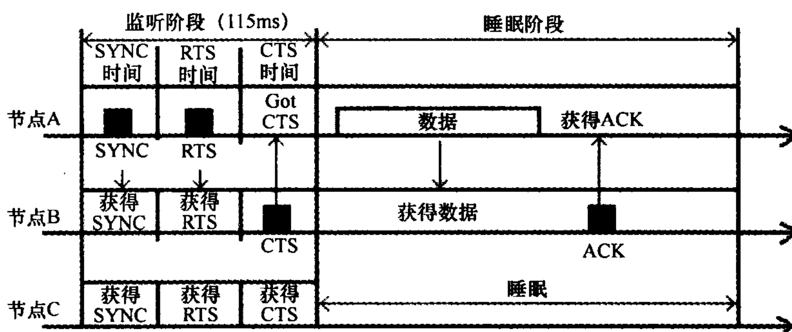


图9-9 S-MAC信息传输

S-MAC协议一个重要的特征是信息传递理论，长信息被分成不同的帧以及在极短时间内被发送。有了这个技术，我们能够通过以媒体访问中的不平等为代价来最小化通信代价，从而节省能量。周期性的休眠可能导致高延迟特别是对于多跳路由算法，因为所有最近的节点都有它们自身的休眠时间表。自适应侦听技术被提出是为了增加休眠延迟，因此增加了总体的延迟。在这个技术中，那些侦听邻居传输的节点在传输的末尾被唤醒一段很短的时间。因此，假如这个节点是下一跳节点，它的邻居能够立即传送数据。传输的末尾被称为RTS/CTS数据包的延迟域。

睡眠调度所带来的最大的好处就是减少了空闲侦听的时间，因此在传感器节点中节约了更多能量。但是这是以延迟为代价的。另外，假如数据包不是以侦听节点为目的，那么被用于改进休眠延迟的自适应侦听可能引起串扰或者空闲侦听。而且，休眠和侦听周期是预定义的常量，在通信量易变的情况下可能降低算法的效率。

9.5.2.2 基于时分的MAC协议

由于降低碰撞和空闲侦听能够节约相当数量的能量，TDMA对于传感器网络MAC协议来说是一个十分吸引人的方案。如图9-10所示，TDMA把信道分成N个间隔。在每一个时隙中，仅仅允许一个节点进行数据传输。N个时隙组成一个帧，周期性地循环。

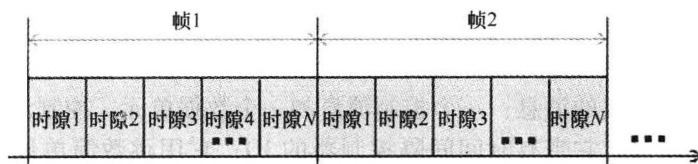


图9-10　TDMA划分时间为有N个时隙的帧

当设计一个基于TDMA的协议时，引起了很多并发现象。由于传感器节点在没有引入大的开销的情况下无法实现大规模协作，从而使时间分配变得困难。必须存在同步功能来纠正由每一个传感器节点的时钟偏移所引起的时间错误。严格的TDMA协议同样需要面对轻量级传输期间所产生的使用问题。

TDMA模式的一些缺点限制了将它应用到WSN中。TDMA通常需要节点来形成簇，类似在细胞通信系统中的细胞。在簇中的一个节点被选择为簇头，且把它作为基站。这种分层结构有几个含义。更重要的是，基于TDMA的协议限制了对于节点数量变化的可扩展性和适应性。当新的节点加入或者旧的节点离开这个簇，基站必须调整帧长度或者时隙分配。另外，帧长度和静态的时隙分配能够限制对于任何给定节点可用的吞吐量，以及在任何簇中活动节点的最大数量。最后，基于TD-MA的协议依靠分布的细粒度的时间同步来排列时隙边界。可能有许多基于基本的TDMA协议的变种，不是调度时隙进行节点通信，可以将时隙分配给带有一些竞争机制的接收者。基站可以动态地一个一个帧的来分配时隙。在Ad hoc设置中，定期地选择节点承担基站的角色，节点轮流扮演基站的角色来平衡能量消耗[49]。总体上，基于TDMA协议能够提供高的能量效率，但是它们不能灵活地适应节点密度或者移动性的改变，并且缺少点对点的通信。

轻量级的MAC

轻量级的MAC（Lightweight MAC，LMAC）协议是基于TDMA的范例。时间被分为帧，每一个帧又被分为时隙，时隙内节点可以传输数据，而不需要竞争媒介或者处理传输碰撞所造成的能量浪费。每一个节点周期性地得到一个时隙，从而允许它控制无线媒介去完成其数据传输。当一个节点有数据需要发送时，它等待其他传输节点的干扰。图9-11说明了LMAC协议的帧格式。

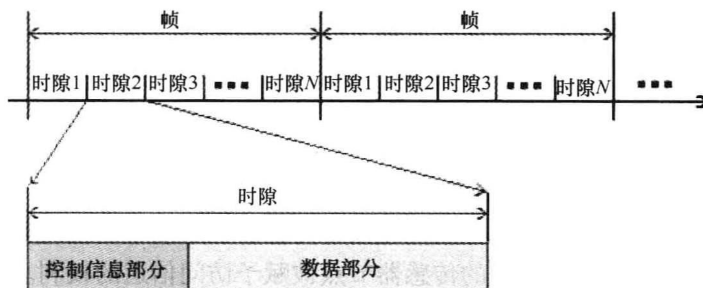


图9-11　LMAC帧格式

不像传统的基于 TDMA 的系统，在 LMAC 协议中的时隙不是通过一个中心管理员在网络节点中进行划分的，而是使用一个分布式算法。在这个时隙中，节点总是发送由两部分组成的消息：一个控制消息和一个数据单元。控制消息有固定的大小并且有多种用途。它带有时间间隔控制器的 ID；它用跳数简单地表示了将数据路由到网关的距离；它描述了预期的接收者，并报告了数据单元的长度。另外，控制信息同样被用于维持节点之间的时间同步，因此节点同时传输帧中它们的时隙的序列号。

所有的邻居节点致力于接收其邻居节点的控制信息。当节点在那个消息没有被标识或者消息没有被标识成为一个识别消息，节点将关闭它们的能量消耗收发器，且只在下次时间间隙被唤醒。假如节点被标识，它将侦听可能无法填满整个剩余时间间隙的数据单元。在信息完成传输之后发射器和接收器都关闭其收发器。短的超时时间间隔确保节点在不受控制的时隙的空闲侦听中不浪费能量。在这个协议中，一个节点只可能在每一帧中传输一个单一的信息。

9.5.2.3 基于预定的 MAC 协议

在传感器网络中，基于时间的媒体访问具有捕获大多数使得能量最优化的时机的潜力。假如介质访问是以共享时间为基准的，由串扰、碰撞、空闲模式和不同状态之间的转换所产生的能量浪费能够降到最低。另外，基于时间的媒体判优能够提高延迟的可预见性和避免由于干扰和缓冲区溢出造成的丢包。然而媒体的调度访问问题是 NP 难的［也就是说，至少和非确定性多项式时间问题（Non-deterministic Polynomial time Problem，NP 问题）一样难］，这使得基于时间的 MAC 模式的可扩展性成为一个主要的关注点。而且，基于时间的分布式媒体判优会引入了过多的代价。另外，在节点中维持时钟同步对于加强调度是很重要的，调度问题对于资源限制的传感器节点是一个非平凡的问题。在这篇文章中提到的大多数基于时间的 MAC 协议侧重于解决这些问题，使用预设数据路由上的预留请求，或者使用简化的启发式方法去解决介质访问调度的复杂性。

能量高效的 TDMA MAC 协议调度

对于使用预定需要来处理基于时间的媒体判优[91]已经进行了探索。有数据需要传输的节点对基站提出数据保留的请求，基站回应一个表明媒体访问期间调度的传输控制信息。未包含于传输控制信息中的节点能够关闭其无线收发机。那些被分配了时隙的节点被基站顺序调度进行数据传输。基站使用能源效率来权衡延迟。尽管最好在连续的时隙内捆绑来自一个节点的所有传输，其他节点的传输将会被延迟。

9.5.2.4 基于优先权的 MAC 协议

使用一个随机函数，通过分配优先级给传感器节点或者链接到目的地来控制无线信道的访问。带有高优先级的传感器节点被赋予访问信道的权利。传感器节点的 ID 和时隙数量提供了随机函数的一个输入，随机函数在两个邻近区域建立优先级。

这种协议的一个例子是一系列由 Bao 和 Garcia-Luna-Aceves[53]提出的协议。

NAMA、LAMA 和 PAMA MAC 协议

节点活跃多路访问（Node Activation Multiple Access，NAMA）协议使用把时间划分为帧的 TDMA 协议；每一个帧细分成节。每一个节进一步被划分为段，而每一段是由大量的时隙所组成的。图 9-12 描述了 NAMA 协议的时分结构。每一个节点选择一个单一的段，选择段来平衡整个信道的使用，以及与其他选择同一段的传感器节点竞争。NAMA 为信令消息保留了每一帧的最后一节，从而允许传感器节点加入网络。每一个传感器节点计算它自身以及邻居节点的优先级，使用这些优先级来决定在传感器节点选择的节内，哪个节点能够访问当前时隙。一个传感器节点在一个节中被分配一个特定的时隙是基于优先级的。假如一个传感器节点在给定的两跳邻居之间，对于给定时隙具有最高的优先级，那么这个传感器节点可以传输数据。假如没有传感器节点的优先级映射到一个时隙，那么有最高优先级的传感器节点可能使用这个时隙。关于 NAMA 协议更加详细的知识请见参考文献［53］。链路活跃多路访问（LAMA）协议对于时隙增加了面向接收机的直接序列扩频（Direct Sequence Spread Spectrum，DSSS）（也就是说，一个接收机选择一个与接收机编码相对应的编码）来激活到传感器目的地节点的链路。每一个传感器节点从有限的伪噪声编码集中得到一个编码。在两跳邻居中最高优先级的传感器节点在每一个时间间隔之内可能通过使用分配到接收机的编码来激活一个链路。成对链路活跃多路访问（Pairwise-link Activation Multiple Access，PAMA）协议基于当前时隙分配优先级给链路，改变链路的编码和优先级来激活链路。假如在所有的链接这两个传感器节点的链路中这条链路具有最高的优先级，在两个传感器节点之间的通信链接会被激活。和 LAMA 相类似，使用 DSSS 允许节点在没有中断的前提下和其他节点通信，协议算法防止了相同编码的碰撞。

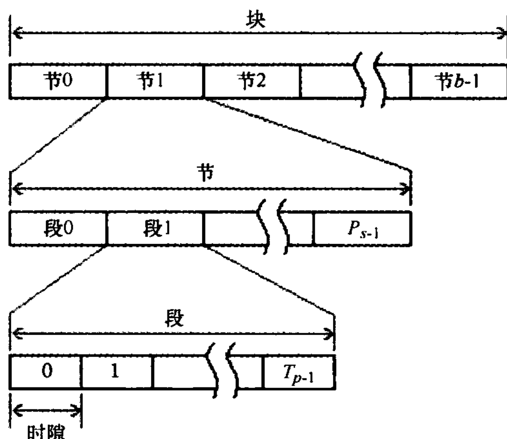


图 9-12 NAMA 协议的时分结构

NAMA、LAMA 和 PAMA 协议的主要不足是计算传感器节点优先级的额外开销，加快了能量消耗并缩短了网络的寿命。另外，LAMA 和 PAMA 需要传感器节点有能控制传播范围的无线电，从而增加了传感器节点的成本。由于优先级基于当前的时隙的数量而变化，动态时隙分配也妨碍了传感器节点形成规则的睡眠调度。

9.5.3 混合 MAC 协议

每一个基于 CSMA、TDMA、FDMA 和 CDMA 的方案对于传感器网络中的 MAC 协议的要求提供了一些优势，同时也带来了一些缺点。混合的 MAC 协议集合了基于竞争协议和预定协议的优势而抵消自身的弱点，能更好地满足这些要求。混合 MAC 协议最大的优势来自于它的简单性以及能够快速适应传输状态，从而可以节约大量能量。在参考文献［55，57，92，93］中提到的协议是这类协议的很好的例子。在这节中，我们简单地讨论这种协议的一个例子。

9.5.3.1 基于前置的混合 MAC 协议

结合了预定和非预定 MAC 协议的优势，基于前置的 MAC 协议可能对能源利用更加高效，并且对于传输类型的变化更加敏感。由前置混合技术产生的能量消耗的降低主要来源于两个方面：第一，当需要时以及仅仅对于传输期间，同步邻近的传感器节点；第二，使用基于前置的协议对于传输模式非常地敏感这一实际情况。从而协议能够基于传输的改变而改变其操作，从而显著地节省能量资源。然而，对于大多数基于前置的协议最重要的缺点是，它们使用长的前置码引起了延迟的增加。在这些形式的协议中长延迟限制了将其应用于实时的或者延迟敏感的应用中。另外，因为这些协议使用混合的预定和非预定协议，协议变得非常的复杂。

无线 MAC 协议（WiseMAC）

WiseMAC 协议是带有前置抽样的混合的 TDMA/CSMA 协议，所有的传感器节点定义了两个通信信道[54]。数据信道是用 TDMA 方法接入的，而控制信道是用 CSMA 方法接入的。WiseMAC 协议使用带有前置抽样的非持续性的 CSMA（NP-CSMA）协议来减少空闲侦听。在前置抽样技术中，一个前置在每个数据包之前提示每个接收的节点。在网络中的所有节点使用一个共同的时间间隙来对介质进行抽样，但是它们相对的调度偏移是独立的。假如节点苏醒后对介质进行抽样发现介质是忙的，它将继续侦听直到其接收一个数据包或者介质重新变得空闲。初步确定的前置的尺寸必须等于采样周期。

为了减少由重新确定固定前置长度而引起的能量消耗，WiseMAC 提供了一个方法来动态地确定前置的长度。这个方法使用了发送节点的直接邻居的睡眠调度知识。另外一个影响激活前置长度的参数是在源和目的地之间潜在的时钟漂移。

WiseMAC 主要的缺点是分散式的休眠-侦听调度时间表导致一个节点的每个邻居具有不同的睡眠和苏醒时间。这对于广播形式的通信是一个特别重要的问题，因为广播的数据包将会缓存在每个睡眠的邻居中，当每个邻居苏醒后会传递很多次。

然而，这种冗余的传输将导致了长延迟和能量消耗。而且，WiseMAC 也会带来隐藏终端问题，这是因为 WiseMAC 是基于非持续性的 CSMA。

9.5.3.2　基于预定的混合协议

结合竞争和分时模式，基于预定的协议的性能能够被提升。PARMAC 协议是基于这种思想来显著地降低能耗的。

基于能量感知预定的 MAC 协议

基于能量感知预定的 MAC 协议（The Power-Aware Reservation-Based MAC, PARMAC）是一个能量感知的协议主要是为 Ad hoc 网络设计的，也适用于传感器网络[92]。这个方法实际上是结合了基于竞争和基于预定的媒体仲裁模式。网络被分为网格，假设每个节点能够到达它所在网格内的所有其他节点。时间被分为固定的帧。为每个网格分配不同的帧。每一个帧是由预定期（Reservation Period，RP）和自由竞争期（Ccontention Free Period，CFP）组成的。在每一个 RP 中，网格单元内的节点交换三条信息来为数据传输、接收以及交换确认保留时隙。数据随后在 CFP 阶段被发送出去。所有节点的时钟都被假设为是同步的。协议允许节点在 CFP 期间休眠来最小化节点的空闲时间从而节省能量。而且，网格间控制数据包代价和包重传是最小的，取得显著的节能效果。然而，网格内竞争仍然有可能存在，并且如果应用需要在不同网格中的节点间进行数据交换这个方法的效率会显著地降低。

9.5.3.3　传输敏感协议

在传感器网络中传输的类型和状态对传感器节点的能量消耗有直接的影响。MAC 协议能够利用这个事实通过自身适应网络状态的变化，显著地节约能量。通信量大的传感器网络为根据通信量调整它们的操作的 MAC 协议提供了一个好的范例。另外，可以利用控制通信和数据通信之间的传输特性的区别，来使得 MAC 协议能够根据传输类型来变换它的操作，这样能够很大程度上降低能量资源消耗。然而，为了实现和执行这些来自于传输特性的益处，MAC 协议应该在传感器网络中跟踪传输特性，并且在传感器网络中的传感器节点共享传输信息。下面是两个例子。

1. 传输自适应媒体访问协议

传输自适应媒体访问（Traffic Adaptive Medium Access，TRAMA）协议试图平衡确定和非确定协议的益处，对于较长数据信息提供不需竞争的预定的时隙和对于较小的周期控制信息提供随机访问时隙[57]。另外，传感器节点通过与邻居分享通信需求和学习它们邻居的两跳拓扑去适应通信和网络状态的变化。TRAMA 协议由三个子协议组成：邻居协议（Neighbor Protocol，NP），用于拓扑信息共享；预定交换协议（Scheduled Exchange Protocol，SEP），允许节点共享那些它们排队的通信和自适应选举算法（Adaptive Election Algorithm，AEA），基于拓扑和通信状态选择用于数据传输的时隙。

TRAMA 协议的帧由多个时隙组成，在帧开始时随机访问控制时隙一起发生，

并且预定数据时隙发生在如图9-13所示的末端。

TRAMA协议的最主要的优势是相对于基于竞争协议获得了更高比例的休眠时间和更低的碰撞率。它主要的缺点是传输间隙比随机存取周期长大约7倍，这意味着占空比很长，因此相对于其他协议，这个协议的能量消耗得更加快。而且，TRAMA协议相对于其他协议具有更高的复杂性，从而限制了其应用。

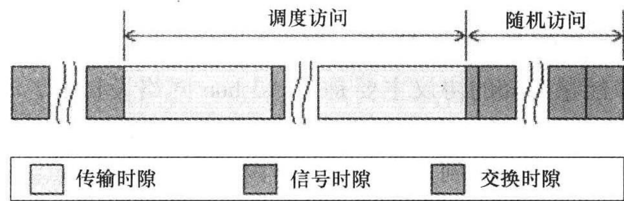


图9-13　TRAMA时隙结构

2. 斑马MAC协议（Zebra MAC Protocol）

斑马MAC（Zebra MAC，Z-MAC）协议使用CSMA作为基础MAC方案，但是使用TDMA时间表作为一个提高竞争解决方案的补充（窍门）[55]。Z-MAC的主要特征是其对网络中的竞争水平的适应性，因此当处于低竞争水平下，它的行为类似于CSMA，而在高竞争水平下，类似于TDMA，从而节约了很大一部分能量。对于经常发生在传感器网络中的动态拓扑改变以及时间同步失败，它的鲁棒性很好。Z-MAC为传感器节点分配TDMA时隙，但是容易允许传感器节点通过带有优先级的后退时间的CSMA去拥有不属于它们自己的时隙。

Z-MAC协议最大的缺点是由TDMA结构引起的大量的消耗。在基于事件的传感器网络中，Z-MAC将需要时间来分发转换到TDMA模式的控制信息，因此延迟就成为了一个问题。在Z-MAC中的延迟问题来自于这个情况，实际上在高竞争周期中Z-MAC使用详细的拥塞通知信息（Explicit Congestion Notification，ECN）去限制隐藏终端的影响。当传感器节点探测到高竞争水平，它传输一个ECN信息给邻居。邻居传播ECN信息到它的邻居节点，从而所有的节点都进入一个高竞争状态（Hign Contention Level，HCL）。在一个时间周期内，假如传感器节点没有进一步地接收到ECN信息，它将回到低水平的竞争状态。

9.5.3.4　基于簇的MAC协议

聚集传感器节点成簇有很多的优点。首先，簇能够区分在本地和全局之间的通信而节约能量。其次，共享本地信息保证了全部状态分布和贪心算法的均衡，全部状态分布对于动态特性的传感器网络会消耗过多的能量，而贪心算法是独立于其他传感器节点对传感器节点自身的行为进行优化的。再次，簇同样使得协议的扩展变得更加容易，因为协议会把簇作为是一个单一的整体。最后，簇可能允许传感器节点去执行某些局部功能，而在全局规模下将消耗更多的能量，例如同步。然而，这

些优势是以协调消息为代价的。管理簇的簇头必须协调传感器节点确保簇平均能量降低。协议经常在传感器节点之间转变簇的作用去平衡由管理操作而引起的额外的能量消耗。节点的动态性使簇协议变得更加复杂，因为簇形成和簇头分配算法必须适应重新部署或者传感器节点死亡。

1. 自适应低能耗结构化分簇协议

自适应低能耗结构化分簇协议（Low-Energy Adaptive Clustering Hierarchy, LEACH）是一种基于分簇的协议，能够在传感器网络中最大限度地减少能量消耗[109]。LEACH 的目的是随机地选择传感器节点作为簇头，因此在通信中基站的高能耗就转移到传感器网络中的所有传感器节点，LEACH 的操作被分为两个阶段，簇建立阶段和稳定阶段。为了尽量地减少开销，稳定阶段持续的时间要比簇建立阶段持续的时间长得多。

在簇建立阶段，传感器节点在 0 和（到）1 之间随机选择一个数。假如这个随机的数字达不到一定的阈值，则传感器节点被选择为簇头。在簇头被选择之后，簇头被传播到网络中的所有传感器节点作为新的簇头。一旦传感器节点接收到了广播信息，它根据来自于簇头的广播信息的信号强度决定自己归属于哪个簇。传感器节点通知适当的簇头，它们将会成为该簇的成员之一。然后，簇头根据 TDMA 方法分配传感器节点能够发送数据到簇头的时隙。

在稳定阶段期间，传感器节点能够开始感知和发送数据到簇头。簇头同样在发送数据到基站之前从其簇中的节点汇总数据。经过一段时间的稳定阶段，网络再次进入簇建立阶段，重新开始选择簇头的另一个循环.

2. 传感器网络高能效的 MAC 协议

传感器网络高能效的 MAC 协议（An Energy-Efficient MAC Protocol for Sensor Networks, GANGS）把传感器节点汇聚成簇[52]。GANGS 协议利用一个未指定竞争协议进行簇内通信和在簇之间基于 TDMA 通信协议进行数据传输。GANGS 不假设传感器节点能够和基站之间进行通信，因此簇头在传感器网络中必须使用单独的路由协议形成一个骨干路由。在 GANGS 中的簇形成由两部分组成：簇头选择和簇连接。由于簇头的附加功能，它执行操作最终将会比邻近的传感器节点具有更低的能量。这时，传感器节点重新执行簇形成的过程，因此能够平衡整个传感器网络中节点的能量消耗。为了分配时隙，簇头执行一个分布式的算法，结果在每一个簇头拥有一个间隙去发送数据，并且知道它的每个邻居所使用的时隙。在簇头决定了 TDMA 时间表之后，它们在簇内部分发消息，从而其他传感器节点能够在最后的帧中使用未分配的时隙来发送它们的数据。

GANGS 协议在簇形成和重建的能量消耗以及需要的时间方面有不足。而且，在 GANGS 中的时隙组织同样存在浪费，因为不是所有的时隙都能够被用到。尽管有这些不足，GANGS 协议为转发通信提供了自由竞争的通信量，同时在簇内保持了随机访问协议的灵活性和简单性。

9.5.4 特定服务质量的MAC协议

延迟、吞吐量和时延抖动的概念在大多数现行的传感器网络中不是最主要的问题。然而，随着对实时应用传感器网络越来越关注，对于协议设计带来了更多的挑战。例如，处于传感器网络监测区域中基于事件触发的实时通信，端到端的延迟必须在一个可接受的范围内，并且这种延迟的变动也必须是可接受的[94]。这样的性能指标通常被称作为通信网络的服务质量（QoS）。因此，收集实时的感知数据要求能量和QoS保证的MAC协议来保证传感器节点的能源有效性以及对收集的测量数据的有效传递。

然而，由于传感器节点严格的资源限制（能量和内存限制），而且必须运行在恶劣的环境下[95]，在传感器网络中获得QoS保障是一项有挑战性的任务。

在WSN中提供QoS保证得到了越来越多研究者的关注。最近在WSN中很多支持一些类型的服务质量的MAC协议已经出现[96-102]。这里有一些例子。

9.5.4.1 传感器网络的QoS控制

参考文献［96］的作者明确地利用节点的冗余。他们对于每个节点开发了一个自适应方案来独立地决定是否进行传输，因此在每一个时隙中都有固定总数的传输会发生。协议允许基站通过广播信道向网络中的每个传感器节点广播QoS信息来完成它的任务，通过使用古尔游戏的数字模式来动态地调整传感器数量到最佳数量，协议使通电的传感器节点和需要关闭的传感器的数量取得平衡，通电的传感器节点用于收集足够的数据来满足QoS要求，而关闭电源则是用来节省相当数量的能量，因此最大化地延长了网络的寿命。这里，QoS的概念是定义为在每个时隙中为了收集足够的数据所发生的总的通信量（即，信息质量）。

9.5.4.2 无线传感器网络协议的一种能量高效的QoS保证MAC协议

Q-MAC（Qos-Aware MAC）方案试图基于优先级区分网络服务来提供QoS，从而在多跳WSN中将能量消耗降低到最小[97]。优先级的水平反映了来自于不同的传感器节点的数据包的重要性。Q-MAC协议通过两个步骤完成其任务：内部节点和外部节点调度。外部节点的调度算法采用多队列结构，根据它们的应用和MAC层抽象而去分类数据包。内部节点调度使用了一个修改后版本的MACAW协议来协调和调度传感器节点之间的数据传输。

9.5.5 跨层的MAC协议

所有目前已经讨论过的MAC协议都是通过利用WSN的协作性质和其相关特性，在一定的程度上增加能量效率。然而，这些协议最主要的共性是它们延续了传统的层次协议结构。这些协议能够在每一个层的某些指标方面获得很高的执行性能，但它们在降低能量消耗时，并不能结合起来最优化网络的整体性能。考虑到WSN的能源稀缺和处理能力有限，共同优化和设计网络层，即跨层设计，对于低

效的传统层协议，结构是非常有前景的。最近，对于 WSN 协议跨层的开发已经取得了重要的成果。事实上，最近的研究成果表明跨层整合以及设计工具对于能量节约有很大的提升。总体上，很多原因形成了这个提升。首先，稀缺的能量、存储器和无线传感器节点的处理能力使这样一种方法成为一种必要。分层协议的显著代价会造成很高的无效性。其次，以近期的经验研究表明在协议设计中需要考虑低能量无线收发器的性能和无线信道的状态。最后，基于事件的无线传感器网络需要感知应用的通信协议。在这一节中，我们回顾了一些最近的为 WSN 设计的通信协议，这些通信协议主要关注于跨层的设计方法。协议根据 OSI 网络栈中不同的层的交互作用分类。在参考文献［31］提出如下分类，其他的协议也有提到。

9.5.5.1 MAC + PHY

在参考文献［65］中，对三个不同的 MAC 协议的物理层（PHY）和 MAC 层的能量消耗进行了分析。作者提供了能量消耗分析并且得出结论，如果使用真实的无线模型单跳通信能够更加的有效。尽管这个结果是很有趣的，但是分析是基于线性网络的基础上的，在真实网络中可能并非如此。

在参考文献［66］中提到了在 MAC 和 PHY 层中的一个跨层的解决方案。这是一个新的基于跨层的载波监听机制用于减轻暴露/隐藏的节点问题，被称为是 MP 方案。这个方案使用了基于 MAC 地址的物理载波感知去决定媒介是否是繁忙的。在 MP 中，一个包的发射器的地址和接收器的地址被纳入到 PHY 头部。在其载波监听操作中使用这个地址信息，节点能够大大地降低暴露/隐藏节点的不利影响。结果表明所提及的方案比先前的方案更有效和更高效。

9.5.5.2 MAC + 网络

MAC 和路由跨层合作对于基于接收器的路由已经在很多的研究文章中被提及[67-69]。在这些文章中作者讨论了能量效率、延迟和算法的多跳性能。参考文献［70］扩展了参考文献［68，69］的工作到一个单一的无线电节点。

在参考文献［71］中，提出了 MACRO 协议，在媒体访问层路由决策是作为成功竞争的结果。更具体的是，下一跳的选择是基于加权进步因素和传输能量相继的增加，直到发现最高效的节点。而且，使用了打开-关闭调度方案。

MAC-CROSS 在参考文献［72］中被提出，MAC-CROSS 协议最小化了完成通信过程所需的节点数。协议利用了路由信息来激活那些仅仅在路由路径上的节点。所有其他的不在路由路径上的节点能够保持休眠模式直到下一个工作周期开始。然而，在能量消耗方面获得的提高是以协议延迟为代价的。

在参考文献［73］中提及一种 WSN 中的针对周期传输的联合时序和路由方案。在这个方案中，节点对于在网络中的每个传输形成了分布式的通断时序，同时路由被建立起来，节点仅仅在需要的时候被激活。由于传输是周期性的，保持这个方案来满足最高的效率。作者同样探索了在通断时序和网络连通性之间的均衡。

在参考文献［74］中同样讨论了在跨层路由和 MAC 框架中使用通断时序。在

这项工作中，设计了基于 TDMA 的 MAC 方案，这里节点是通过局部的拓扑信息分布式地选择它们适当的时隙。路由协议对于路由的确定同样使用了这类信息。

所有这些方案的性能评价在路由和 MAC 层上呈现了跨层方法的优点。

多跳基础网络体系结构（Multi-hop Infrastructure Network Architecture，MINA）是汇集 MAC 和路由协议的另外一种工作[60]。Ding 等人提出了一个多跳网络体系结构，到基站有相同跳的网络节点被组织在相同的层里。信道访问是结合了 CDMA 或者 FDMA 的基于 TDMA 的 MAC 协议。超帧是由控制数据包、信标帧和数据传输帧组成的。信标和数据帧是时隙的。在分簇网络结构中，所有的簇成员在信标时隙里提交它们的传输请求。相应地，簇头报告了数据帧的时隙。

路由协议是一个简单的多跳协议，每个节点在离基站最近的层里都有一个转发节点。转发节点是基于剩余能量来选择的。Ding 等人将信道分配问题作为一个 NP 完全问题，提出一个次优方案。而且，传感器节点的传输范围是个决策变量，因为它影响到了网络的分层（跳数变化）。运用模拟对于具体实际情况去寻找一个良好的范围值。

9.5.5.3 网络+PHY

参考文献［75］提出了多跳无线网络的一个跨层的网络吞吐量最佳的方案。作者把吞吐量优化问题分成两个子问题，即在网络层中的多跳流路由和在物理层的功率分配。吞吐量与每个链路数据流的比率有关，而数据流比率依赖于链路容量，也就是与节点的功率发射水平有关。另一方面，能量分配问题和干扰以及链接的速率是联系在一起的。基于这个解决方案，提出了一个基于 CDMA/OFDM 的解决方案，功率控制和路由都以分布式方式执行。

在参考文献［76］中，针对地理位置路由提出了新的转发策略。作者为具有自动重复请求（ARQ）以及没有 ARQ 的最佳转发路由的网络提供了表达式。而且，提供了这些情况的两种转发策略。转发算法需要每个邻居数据包的接收率来决定下一跳及建立相应的路由建立。尽管新的转发指标表明了在 WSN 中的跨层转发技术的优势，分布式最佳一跳距离的分析是基于线性网络结构的。

9.5.5.4 传输+PHY

在参考文献［77］中，考虑了功率控制和拥塞控制的跨层最优化解决方案。作者提供了功率控制和拥塞控制的理论分析，以及分层和跨层之间的平衡。基于这个结构体系，提出了一个基于 CDMA 的跨层通信协议，能够控制传输功率和传输速率。然而，这个解决方案仅仅适用于基于 CDMA 的无线多跳网络，对于 CDMA 技术不可行的 WSN 来说不可用。

9.5.5.5 三层解决方案

除了被提及的关注于成对的跨层交互协议以外，在三层协议层间存在更多一般的跨层方法。在参考文献［78］中，提出了最优化的传输功率，传输速率以及基于 TDMA 的链接调度。最优化是指能够最大化网络寿命，而不是最小化总的平均功率消耗。

在参考文献［79］中，提出了通过联合路由、时间表和链接层最优化来最大化网络寿命的自适应策略。作者提出了一个可变长度的 TDMA 方案，时隙长度是根据路由需要最优化分配的，并最小化了网络中的能量消耗。最优化问题所考虑的能量消耗包括传输能耗和路由能耗。基于这个分析，研究表明单跳通信在电路能耗而不是传输能耗占能量消耗支配地位的情况下可能是最佳的。尽管在文章中提到的最优化问题是很有深度的，但并未提出实际可行的通信协议。而且，没有考虑传输层的问题，例如拥塞和流量控制。

9.6　IEEE802.15.4/ZigBee MAC 协议

ZigBee[104]是一种新的无线技术，根据 IEEE802.15.4 个域网网络标准而建立起来的[105]。它主要设计用于广泛的自动化运用和代替现有的非标准技术。目前在欧洲它是工作在 868MHz 的带宽，采用 20kbit/s 的数据速率；在美国它是运行在 914MHz 的带宽，采用 40kbit/s 的数据速率；以及 2.4GHz 的 ISM 全球频段，最大 250kbit/s 的数据速率。

ZigBee 被设计为设备提供低成本和低功耗的连接，要求电池寿命达到从几个月到几年时间，但是不需要高的数据传输速率。它的一些主要特征是基于标准的无线技术、低数据传输速率、低能量消耗、简单的和低成本的无线网络等。有时，人们混淆了 IEEE 802.15.4 和 ZigBee，ZigBee 是 ZigBee 联盟的一个新标准。ZigBee 使用由 IEEE 802.15.4 提供的服务，并且增加了网络构造（星形网络、对等网/网状网络、分簇-树状网络）、安全性、应用服务以及更多。IEEE 802.15.4 MAC 协议是一个混合的协议，结合了基于时隙和基于竞争的方案。在这节中我们简单地描述了 IEEE 802.15.4/ZigBee 标准，重点介绍了 MAC 层。

9.6.1　IEEE 802.15.4/ZigBee 协议栈架构

参照继承了标准 OSI 模型，ZigBee 协议栈如图 9-14 所示的层次结构。在最开始的两层中，物理层（PHY）和媒体访问层（MAC）是由标准的 IEEE 802.15.4 定义的[105]。其上面的层的是由 ZigBee 设计的[104]。

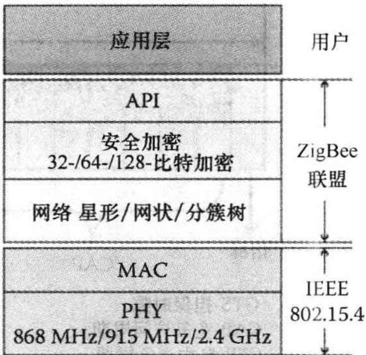


图 9-14　IEEE 802.15.4/ZigBee 协议栈架构

9.6.2　ZigBee 网络架构

在 ZigBee 网络中有两种不同类型的设备类型：全能设备（Full-Function Device，FFD）和降低功能设备（Reduced-Function Device，RFD）。作为一个个域网（Personal Area Network，PAN）协调

器，一个协调器；或者一个设备，FFD能够在三种模式下运行。一个FFD能够和一个RFD或者其他的FFD通话，而一个RFD只能与一个FFD通话。一个遵守IEEE 802.15.4的系统是由很多部件组成的。最基本的模式是设备。一个设备能够被设计为RFD或者FFD。在一个个人操作空间（Personal Operating Space，POS），两个或者更多的设备在相同的物理信道通信组成一个无线个域网（Wireless Personal Area Network，WPAN）。然而，一个网络应该包括至少一个FFD，作为PAN协调器（进行操作）。根据应用需求，小范围的无线个域网（Low Range Wireless Personal Area Network，LR-WPAN）可能运行在以下的两种拓扑结构中：星形拓扑或者点对点拓扑。在星形拓扑结构中，所有的通信和资源预留发生在PAN协调器中。在点对点的拓扑中，设备之间直接的通信不需要通过PAN协调器，但是它们必须提前与PAN协调器联系来加入到网络中去。点对点拓扑允许形成更多复杂的网络构造，例如网状网络拓扑。产业化的控制和监视、WSN，资产和库存跟踪、智能农业以及安全等应用都有可能受益于这种网络拓扑。一个点对点的网络能够成为Ad hoc网络，自组织和自修复。而且同样允许在网络中从任何一个设备到其他任何设备的多跳来路由信息。这种功能能够被增加到网络层。IEEE 802.15.4关注于星形拓扑网络，但是留下了很多的点对点网络的选项和功能没有定义。因此，在本节的其余部分，我们集中讨论在星形网络中协调器和设备之间的数据交换。

9.6.3 超帧结构

星形网络的协调器运行在信标使能模式下，在超帧的帮助下组织信道访问和数据传输。超帧典型地使用于低延迟设备的环境，尽管在长时间内不活跃，但这种连接必须保持下去。超帧的结构如图9-15所示。所有的超帧的长度都是相同的。协调器通过发送一个信标帧数据包而开始每一个超帧。信标帧包括一个超帧，详细地描述了下面的超帧的各个组成部分的长度。

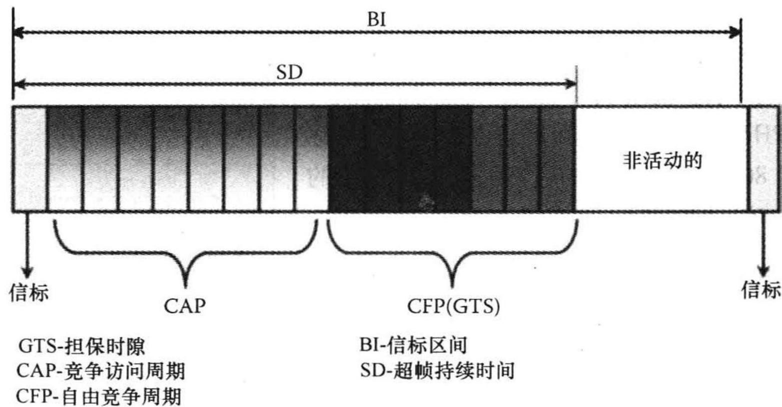


图9-15　IEEE 802.15.4 超帧结构

超帧细分为活跃期和非活跃期。在非活跃期期间，所有的节点包括协调器能够关闭它们的接收器，进入休眠状态。节点能够在非活跃期结束之前立即激活来接收下一个信标。

活跃期被细分为 16 个时隙。第一个时隙是被信标帧所占用，而剩余的时隙被划分为一个竞争访问周期（Contention Access Period，CAP）以及很多的连续担保时隙（Guaranteed Time slot，GTS），应用能够使用一个或者更多时隙来传输数据。

活跃周期的长度、非活跃周期以及一个单一的时隙和 GTS 时隙的使用的长度是可以调节的。

在整个活跃周期中协调器是激活的。在 GTS 期间仅仅在为它们分配时隙后，相关设备才处于激活状态；而在其他的 GTS 时隙中，它们会进入休眠模式。在 CAP 中，假如设备自身没有任何数据要传输或者从协调器中获取数据，它能够关闭它的接收器。

9.6.4　数据传输

有三种不同类型的可能的数据传输。它们是：从设备传输到协调器，从协调器重传输到设备，以及最后在任何两个设备之间的传输。在一个星形网络拓扑中仅仅只有前面的两个传输技术是可能的。虽然不支持任何两个设备之间的传输，但在点对点的网络拓扑中所有三种类型的传输都是可能的。传输能够重新在这两种模式中的任意一种模式中执行，这主要依赖于信标传输是否是使能的。我们集中讨论信标使能的网络。超帧的信标周期必须足以给网络结构和它的设备提供服务。数据传输到协调器需要一个信标同步阶段，假如合适，接下来就是 CSMA/CA 传输（如果使用超帧，那么以时隙的方式），确认是可选的。来自协调器的数据传输通常按照设备的要求：假如信标是可用的，这些都是用来发送请求信号；协调器确认了请求随即发送一个设备能够识别的数据包。在超帧不可用的情况下使用同样的方法，只有在这种情况下没有信标跟踪未解决的信息。

点对点的网络有可能使用无时隙的 CSMA/CA 或者同步机制，在这种情况下，在任何两个设备之间的通信是可能的，而在结构化模式中其中的一个设备必须是网络的协调器。总体上，所有的执行程序遵循一个请求-确认/指示-响应分类。图 9-16和图 9-17 显示了在 IEEE 802. 15. 4 中的数据传输过程。

尽管 IEEE802. 15. 4 与传感器网络相似，关注于应用，但是它存在很多缺点限制了其在传感器网络中的使用。首先，仅仅在星形拓扑的通信机制中定义了标准，而没有明确地定义点对点的拓扑操作。除此之外，标准不允许不同的协调器之间的互相操作。

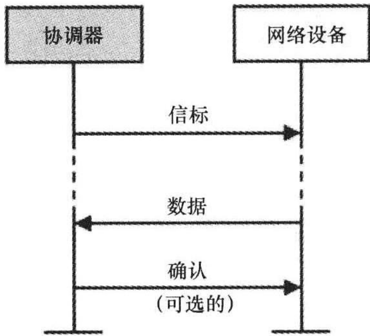


图9-16　IEEE 802.15.4数据传输：数据传输到协调器

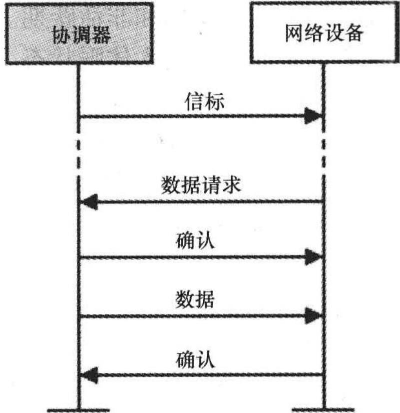


图9-17　IEEE 802.15.4数据传输：来自协调器的数据传输

9.6.5　蓝牙

蓝牙系统是设计作为一个WPAN的主要应用，将其设备连接到个人电脑。它已经被用来作为原型无线网络应用的一种方式。物理层是基于跳频（Frequency Hopping Spread Spectrum，FHSS）方案，它具有1.6kHz的跳频以及合适的跳频序列分配方案。这些节点被组织成微网，微网具有一个主节点以及最多七个主动的从属节点。主节点选择跳频序列，从属节点必须遵循这些调频序列。而且，在一个微网中有很多被动的从属节点。主节点不断查询活跃的从属节点。蓝牙的两个主要的缺点是需要一直有一个主节点，花费很多的能量来激活从属节点，而且不会限制在每个网格中从属节点的数量。当需要很大数量的主节点时，这种情况和密集部署的WSN是不兼容的。一个活动的从属节点必须经常是打开的，因为它不能够预测自身什么时候将会被主节点查询。一个被动的从属节点向主节点申请变成主动的从属节点。假如已经有七个主动的从属节点，申请会失败。而且，要求每个节点能够接替主节点或者从属节点，承担一定的复杂度。同样，在一个微网中，节点需要紧密地同步才能够完成高频跳动操作[110]。

9.7　开放的研究方向

尽管很多的媒体访问方案在传感器网络中被提出，MAC层协议设计仍然需要进行大量的研究。在这节中，我们描述了这些未解决的问题，并提出未来的研究方向。

（1）标准化　在传感器网络中缺少一个MAC协议标准。到目前为止没有协议被认同为标准。主要的原因是MAC协议的选择，总体上，是取决于应用，这就意

味着对于传感器网络的 MAC 有很多的标准。另外一个原因是在底层（物理层）和物理传感器硬件上缺少标准。

TDMA 在自由碰撞媒体访问中有天然的优势。然而它包含了时钟漂移问题，由于空闲间隙，低传输负载时也降低了吞吐量。TDMA 系统的困难在于节点的同步和适应新节点的加入，电池能量的耗尽，由于干扰造成的链接断开，中继节点的睡眠调度，以及由分簇算法引起的调度。因此，时隙分配应该考虑这种可能性。然而，在一个分布式环境下的传统的 TDMA 中改变时隙分配是不容易的，因为所有的节点必须与时隙分配相一致。

FDMA 对于自由碰撞媒体的另外一种方案，尽管在不同的无线信道中动态的通信带来了额外的电路需求，这和客观存在的传感器网络系统相反，增加了传感器节点的成本。

CDMA 同样提供自由碰撞的媒体，但是其高的计算要求对于外在的传感器节点的低能量消耗是一大障碍。如果结果表明 CDMA 的高计算复杂度能够和其防撞功能相均衡，那么可以认为 CDMA 协议对于传感器网络同样是一个好的解决方案。

基于 CSMA 的协议具有较低的延迟，并且保证了低通信负载潜在的吞吐量，这种情况通常发生在 WSN 中。然而，应该采用附加的碰撞避免或者碰撞检测方法来控制可能的碰撞。缺少与在通常的结构体系中的 TDMA、CSMA 或者其他的媒体访问协议之间的对比。

迄今为止，总体上传感器网络和 MAC 最主要的设计目标是能源效率。然而，随着传感器网络新的应用的出现，其他的优化准则（或者 QoS 参数），例如延迟和遵守实时约束，或者可靠的数据传输可能得到重视。一个重要的问题是很多应用需要对多种互相冲突的标准进行优化。因此，应用需要一种方式去实现这些互相冲突的目标之间的均衡。

（2）移动性管理　在传感器网络设计 MAC 协议中考虑移动性在很长一段时间内被认为是一个充满挑战的研究方向，然而即使是在文章中出现的最新的 MAC 协议也不能够明确地在 MAC 层中考虑移动性问题，除了极少数的，例如参考文献[48，63，106]。最近，对医疗保健方面和传感器节点在灾难响应方面的应用引起了人们极大的兴趣，而在这些环境中使用移动的传感器节点。因此，在这一领域有很大的研究空间。

鉴于未来部署大型的微型传感器和在很多的应用方面使用范围的扩大，使得可扩展性问题成为在未来可能的研究方向中最值得注意的问题。

对通信和拓扑改变的适应性：为了节约更多的能量，MAC 协议应该适应网络拓扑和传输特性的改变。然而适应改变通常会增加协议的复杂性，这样有可能带来不利因素，消耗传感器节点的能量，因此仍然需要更多的研究去解决适应性问题。

（3）QoS 处理　尽管上面已经讨论了一些方法，WSN 中的 QoS 领域仍然是一个未开发的领域。这里的问题是如何达到应用的 QoS 需求和能量限制之间的

均衡[29]。

设计协议中的混合方法是一个很有希望的方法。尽管在这方面已经做了大量的工作，针对上面讨论的无线传感器 MAC 协议要求仍然需要进行广泛的研究。

（4）跨层交互　尽管对基于跨层相互作用的新通信协议已经做了大量的研究，但是以一个统一的方式重新考虑网络层的协议功能，从而在一个单一的通信单元来进行无线传感器网络中的有效通信能够获得很多。针对 WSN 协议的跨层设计系统技术的发展，有几个开放的研究问题。更加详细的内容见参考文献［31］。

（5）真实的仿真模型　设想在大多数仿真环境下没有必要反映真实世界的情况（例如，一个无线发射区域是圆形的，所有的无线电有相等的通信范围，等等）。为了完全地理解设计 MAC 协议的复杂性和开发在实际情况下的方案，有必要开发更加实际的无线电和能量仿真模型。重新认识 Kotz 关于无线网络研究的错误的规律（参考文献［108］），从而理解为什么 MAC 协议在仿真时产生极其精确的结果而在实际部署情况下却是完全不同的，这是非常重要的。

为了对 MAC 层性能获得更多的现实依据，传感器网络研究者应该从仿真转变为到原型或者实际环境下的试验。大体上，考虑到 MAC 性能现有的无线 MAC 协议集中讨论能量最优化系统，仍然没有充分地考虑传感器网络的所有需求。最重要的挑战仍然是提供可预见的延迟或优化保证，同时最小化数据包的代价和能量消耗。

9.8 结论

最近几年，WSN 中的 MAC 协议吸引了很多的关注，相对其他的无线网络中其他传统的 MAC 具有独特的挑战。在这章中，我们讨论了 WSN 中 MAC 协议的特殊需求；对现有研究的 WSN 中 MAC 协议进行了分类，提到了很多已经提出的协议；讨论并确定了开放性的问题；以及提出了一系列未来的研究方向。如在上一节所讨论的，尽管很多的 MAC 方案已经在 WSN 中被提及，该领域在很大的程度上仍然是开放研究，如前面所说，没有明确的未来应该努力研究的方向。这具有这极大的吸引力，并且有很多的悬而未决的问题，“是否存在支持各种应用和操作环境的普遍可用的和灵活，而且能够最小化传输能耗同时提供可接受的传输特征的 MAC 协议？”然而，随着目前在 WSN 中的 MAC 协议的研究，我们鼓励更加深入地研究问题和如本章所述地对这些开放性问题的解决策略进行更加深入地研究。

参考文献

1. I. F. Akyildiz et al., Wireless sensor networks: A survey, *Computer Networks*, 38, 393–422, March 2002.
2. R. Min et al., Low power wireless sensor networks, in *Proceedings of International Conference on VLSI Design*, Bangalore, India, January 2001.

3. S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, A taxonomy of wireless microsensor network models, in *ACM SIGMOBILE Mobile Computing and Communications Review*, 6(2), 28–36, April 2002, ACM, New York.
4. J. M. Rabaey et al., PicoRadio supports ad-hoc ultra low power wireless networking, in *IEEE Computer*, 33, 42–48, July 2000.
5. G. J. Pottie and W. J. Kaiser, Wireless integrated network sensors, *Communication of the ACM*, 43(5), 51–58, 2000.
6. R. H. Katz, J. M. Kahn, and K. S. J. Pister, Mobile networking for smart dust, in *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, MobiCom'99*, Seattle, WA, August 1999.
7. T. Bokareva, W. Hu, S. Kanhere, B. Ristic, N. Gordon, T. Bessell, M. Rutten, and S. Jha, Wireless sensor networks for battlefield surveillance, in *Proceedings of the Land Warfare Conference, LWC-2006*, Brisbane, Queensland, Australia, October 2006.
8. A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, Wireless sensor networks for habitat monitoring, in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 88–97, September 2002, Atlanta, GA.
9. G. Tolle, J. Polastre, R. Szewczyk, N. Turner, K. Tu, S. Burgess, D. Gay, P. Buonadonna, W. Hong, T. Dawason, and D. Culler, A macroscope in the redwoods, in *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems (SenSys'05)*, pp. 51–63, November 2005, San Diego, CA.
10. N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, A wireless sensor network for structural monitoring, in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys'04)*, pp. 13–24, November 2004, Baltimore, MD.
11. K. Srinivasan, M. Ndoh, H. Nie, H. Xia, K. Kaluri, and D. Ingraham, Wireless technologies for condition-based maintenance (CBM) in petroleum plants, in *Proceedings of the 1st International Conference on Distributed Computing in Sensor Systems (DCOSS'05)*, Poster Session, June 30–July 1, 2005, Marina del Rey, CA.
12. K. Sohrabi et al., Protocols for self-organization of a wireless sensor network, *IEEE Personal Communications*, 7(5), 16–27, October 2000.
13. A.L. Buczak, V.R. Jamalabad, Self-organization of a heterogeneous sensor network by genetic algorithms, in *Intelligent Engineering Systems through Artificial Neural Networks*, eds., C. H. Dagli, M. Akay, A. L. Buczak, O. Ersoy, B. R. Fernandez, vol. 8, ASME Press, New York, 1998, pp. 259–264.
14. N. Abramson, The ALOHA system—another alternative for computer communications, in *Proceedings of AFIPS Conference*, vol. 36, pp. 295–298, 1970.
15. L. Kleinrock and F. Tobagi, Packet switching in radio channels: Part I -carrier sense multiple access modes and their throughput-delay characteristics, *IEEE Transactions on Communications*, 23(12), 1400–1416, December 1975.
16. L. Roberts, ALOHA packet system with and without slots and capture, Stanford Research Institute, Advanced Research Projects Agency, Network Information Center, Tech. Rep. ASS Note 8, 1972.
17. I. Demirkol, C. Ersoy, and F. Alagoz, MAC protocols for wireless sensor networks: A survey, *IEEE Communications Magazine*, 44(4), 115–121, April 2006.

18. W. Ye and J. Heidemann, Medium access control in wireless sensor networks, USC/ISI Technical Report ISI-TR-580, October 2003.
19. K. Kredo II and P. Mohapatra, Medium access control in wireless sensor networks, *Journal of Computer Networks*, 51, 961–994, 2007.
20. A. Bohm, State of the art on energy-efficient and latency constrained networking protocols for wireless sensor networks, HH, Technical Report, IDE0749, June 2007.
21. K. Langendoen, Meduim access control in wireless sensor networks, Book chapter, *Practice and Standards*, vol. II, eds., H. Wu and Y. Pan, 2007.
22. P. Baronti, P. Pillai, V. W. C. Chook, S. Chessa, A. Gotta, and Y. Fun Hu, Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards, *Journal of Computer Communication*, 30, 1655–1695, 2007.
23. P. P. Czapski, A survey: MAC protocols for applications of wireless sensor networks, in *Proceedings of IEEE TENCON 2006, A Technical Conference of the IEEE Region*, vol. 10, pp. 1–4, Hong Kong, November 14–17, 2006.
24. K. Langendoen and G. Halkes, Energy-efficient medium access control, Book chapter in the *Embedded Systems Handbook*, ed., R. Zurawski, CRC Press, ISBN: 9780849328244, August 2005.
25. Jurdak, R., C. V. Lopes, and P. Baldi, A survey, classification and comparative analysis of medium access control protocols for ad hoc networks, in *IEEE Communications Surveys*, 2004.
26. P. Naik and K. M. Sivalingam, A survey of MAC protocols for sensor networks, Book chapter, *Wireless Sensor Networks*, Kluwer Academic Publishers, Norwell, MA, pp. 93–107, 2004.
27. D. Ganesan, A. Cerpa, W. Ye, Y. Yu, J. Zhao, and D. Estrin, Networking issues in wireless sensor networks, *Journal on Parallel and Distributed Computing*, 64, 2004.
28. J. N. Al-Karaki and A. E. Kamal, Routing mechniques in wireless sensor networks: A survey, *IEEE Journal on Wireless Communications*, 11(6), 6–28, December 2004.
29. D. Chen and P. K. Varshney, QoS support in wireless sensor networks: A survey, in *Proceedings of the International Conference on Wireless Networks 2004, ICWN '04*, Las Vegas, NV, June 2004.
30. M. Ali, U. Saif, A. Dunkels, T. Voigt, K. Rmer, K. Langendoen, J. Polastre, and Z. A. Uzmi, Medium access control issues in sensor networks, *ACM SIGCOMM Computer Communication Review*, 36(2), 33–36, April 2006.
31. T. Melodia, M. C. Vuran, and D. Pompili, The state of the art in cross-layer design for wireless sensor networks, *Network Architect. 2005*, LNCS 3883, pp. 78–92, Springer-Verlag, Berlin, Heidelberg, 2006.
32. W. Ye, J. Heidemann, and D. Estrin, Medium access control with coordinated adaptive sleeping for wireless sensor networks, *IEEE/ACM Transactions on Networking*, 12(3), 493–506, June 2004.
33. W. Ye, J. Heidemann, and D. Estrin, An energy-efficient MAC protocol for wireless sensor networks, in *Proceedings of IEEE INFOCOM*, New York, June 2002.
34. C. Jones, K. Sivalingam, P. Agrawal, and J. C. Chen, A survey of energy efficient network protocols, *Wireless Networks*, 7, 2001.

35. P. Lettieri and B. Srivastava, Advances in wireless terminals (I), *IEEE Personal Communications Magazine*, 6(1), 6–19, February 1999.
36. P. Havinga, G. Smit, and M. Bos, Energy efficient wireless ATM design, *Mobile Networks and Applications*, 5(2), 147–155, Kluwer Academic Publishers, Hingham, MA, June 2000.
37. M. Miller and N. Vaidya, Minimizing energy consumption in sensor networks using a wakeup radio, in *Proceedings of the IEEE International Conference on Wireless Communications and Networks, WCNC'04*, Atlanta, GA, March 2004.
38. E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. A. Wang, and A. Chandrakasan, Physical layer driven protocol and algorithm design for energy efficient wireless sensor networks, in *Proceedings of the 7th ACM/IEEE Conference on Mobile Computing and Networks MOBICOM'01*, Rome, Italy, July 2001.
39. L. Kleinrock and F. Tobagi, Packet switching in radio channels: Part II—The hidden terminal problem in CSMA and busy tone solutions, *IEEE Transactions on Communications*, 23(12), 1417–1433, 1975.
40. S. Singh and C. Raghavendra, PAMAS—Power aware multi-access protocol with signaling for ad hoc networks, *SIGCOMM Computer Communications Review*, 28(3), 5–26, 1998.
41. M. C. Vuran and I. F. Akyildiz, Spatial correlation-based collaborative medium access control in wireless sensor networks, *IEEE/ACM Transactions on Networking*, 14(2), 316–329, IEEE Press, Piscataway, NJ, April 2006.
42. K. Jamieson, H. Balakrishnan, and Y. C. Tay, Sift: A MAC protocol for event-driven wireless sensor networks, MIT Laboratory for Computer Science, Tech. Rep. 894, May 2003.
43. J. Polastre, J. Hill, and D. Culler, Versatile low power media access for wireless sensor networks, in *Proceedings of the International Conference on Embedded Networked Sensor Systems, SenSys'04*, pp. 95–107, 2004.
44. E.-Y. A. Lin, J. M. Rabaey, and A. Wolisz, Power-efficient Rendez–Vous schemes for dense wireless sensor networks, in *Proceedings of the IEEE International Conference on Communications, ICC'04*, vol. 7, pp. 3769–3776, 2004.
45. G. Lu, B. Krishnamachari, and C. S. Raghavendra, An adaptive energy efficient and low-latency MAC for data gathering in wireless sensor networks, in *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS'04)*, pp. 224, April 26–30, 2004, Santa Fe, NM.
46. T. V. Dam and K. Langendoen, An adaptive energy-efficient MAC protocol for wireless sensor networks, *The First ACM Conference on Embedded Networked Sensor Systems, Sensys'03*, Los Angeles, CA, November, 2003.
47. P. Lin, C. Qiao, and X. Wang, Medium access control with a dynamic duty cycle for sensor networks, in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'04)*, vol. 3, pp. 1534–1539, March 21–25, 2004, Atlanta, GA.
48. H. Pham and S. Jha, An adaptive mobility aware MAC protocol for sensor networks, in *The 1st IEEE International Conference on Mobile Ad Hoc and Sensor Networks*, October 24–27, Tampa, FL, 2004.
49. Y. E. Sagduyu and A. Ephremides, The problem of medium access control in wireless sensor networks, *IEEE Wireless Communications*, 11(6), 44–53, 2004.

50. L. van Hoesel and P. Havinga, A lightweight medium access protocol (LMAC) for wireless sensor networks, in *1st Int. Workshop on Networked Sensing Systems INSS'04*, Tokyo, Japan, June 2004.
51. L. van Hoesel and P. Havinga, Poster abstract: A TDMA-based MAC protocol for WSNs, in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys'04)*, pp. 303–304, 2004, Baltimore, MD.
52. S. Biaz and Y. D. Barowski, GANGS: An energy efficient MAC protocol for sensor networks, in *Proceedings of the 42nd Annual Southeast Regional Conference*, pp. 82–87, 2004, Huntsville, AL.
53. L. Bao and J. Garcia-Luna-Aceves, A new approach to channel access scheduling for ad hoc networks, in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking, MobiCom'01*, pp. 210–221, July 2001, Rome, Italy.
54. A. El-Hoiydi, Spatial TDMA and CSMA with preamble sampling for low power ad hoc wireless sensor networks, in *Proceedings of ISCC 2002, The Seventh International Symposium on Computers and Communications*, pp. 685–692, July 1–4, 2002.
55. I. Rhee, A. Warrier, M. Aia, and J. Min, Z-MAC: A hybrid MAC for wireless sensor networks, Technical report, Department of Computer Science, North Carolina State University, Raleigh, NC, April 2005.
56. I. Rhee, A. Warrier, and L. Xu, Randomized dining philosophers to TDMA scheduling in wireless sensor networks, Technical report, Computer Science Department, North Carolina State University, Raleigh, NC, 2004.
57. V. Rajendran, K. Obraczka, and J. Garcia-Luna-Aceves, Energy-efficient collision free medium access control for wireless sensor networks, in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems, SenSys'03*, pp. 181–192, November 2003, Los Angeles, CA.
58. V. Ekanayake, C. Kelly, and R. Manohar, An ultra low-power processor for sensor networks, *ACM SIGPLAN Notices*, 29(11), 27–36, November 2004.
59. S. Cui, R. Madan, A. J. Goldsmith, and S. Lall, Joint routing, MAC, and link layer optimization in sensor networks with energy constraints, in *Proceedings of the IEEE International Conference on Communications, ICC'05*, vol. 2, pp. 725–729, Korea, May 16–20, 2005.
60. J. Ding, K. Sivalingam, R. Kashyapa, and L. J. Chuan, A multi-layered architecture and protocols for large-scale wireless sensor networks, in *IEEE 58th Vehicular Technology Conference, VTC'03*, vol. 3, pp. 1443–1447, October 6–9, 2003, Orlando, FL.
61. M. Zorzi, A new contention-based MAC protocol for geographic forwarding in ad hoc and sensor networks, in *IEEE International Conference on Communications (ICC'04)*, vol. 6, pp. 3481–3485, June 20–24, 2004, Paris, France.
62. R. Rugin and G. Mazzini, A simple and efficient MAC-routing integrated algorithm for sensor network, in *IEEE International Conference on Communications (ICC'04)*, vol. 6, pp. 3499–3503, June 20–24, 2004, Paris, France.
63. M. Ali, T. Suleman, and Z. A. Uzmi, MMAC: A mobility-adaptive, collision-free MAC protocol for wireless sensor networks, in *Proceedings of the 24th IEEE IPCCC'05*, Phoenix, AZ, April 2005.

64. J. Ai, J. Kong, and D. Turgut, An adaptive coordinated medium access control for wireless sensor networks, in *Proceedings of the 9th International Symposium on Computers and Communications (ISCC'04)*, vol. 1, pp. 214–219, June 28–July 1, 2004, Alexandria, Egypt.
65. J. Haapola, Z. Shelby, C. Pomalaza-Raez, and P. Mahonen, Cross-layer energy analysis of multi-hop wireless sensor networks, in *Proceedings of the 2nd European Workshop in Wireless Sensor Networks (EWSN' 05)*, pp. 33–44, January 31–February 2, 2005, Istanbul, Turkey.
66. A. Chan and S. Chang Liew, Merit of PHY-MAC cross-layer carrier sensing: A MAC-address-based physical carrier sensing scheme for solving hidden-node and exposed-node problems in large-scale Wi-Fi networks, in *Proceedings of the 31st IEEE Conference on Local Computer Networks, LCN'06*, pp. 871–878, November 14–16, 2006, Tampa, FL.
67. P. Skraba, H. Aghajan, and A. Bahai, Cross-layer optimization for high density sensor networks: Distributed passive routing decisions, in *Proceedings of Ad-Hoc Now04*, Vancouver, British Columbia, Canada, July 2004.
68. M. Zorzi and R. Rao, Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance, *IEEE Trans. Mobile Computing*, 2(4), 337–348, October–December 2003.
69. M. Zorzi and R. Rao, Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance, *IEEE Trans. Mobile Computing*, 2(4), 349–365, October–December 2003.
70. M. Zorzi, A new contention-based MAC protocol for geographic forwarding in ad hoc and sensor networks, in *Proceedings of the IEEE International Conference on Communications (ICC'04)*, vol. 6, pp. 3481–3485, June 20–24, 2004, Paris, France.
71. D. Ferrara et al., MACRO: An integrated MAC/routing protocol for geographical forwarding in wireless sensor networks, in *Proceedings of IEEE INFOCOM'05*, vol. 3, pp. 1770–1781, March 13–17, 2005, Miami, FL.
72. C. Suh, Y. Ko, and D. Son, An energy efficient cross-layer MAC protocol for wireless sensor networks, *LNCS Book Series: Advanced Web and Network Technologies, and Applications Book*, ISBN:978-3-540-31158-4, Springer, Berlin/Heidelberg.
73. M. L. Sichitiu, Cross-layer scheduling for power efficiency in wireless sensor networks, in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'04)*, vol. 3, pp. 1740–1750, March 7–11, 2004, Hong Kong.
74. L. van Hoesel, T. Nieberg, J. Wu, and P. J. M. Havinga, Prolonging the lifetime of wireless sensor networks by cross-layer interaction, *IEEE Wireless Communications*, 11(6), 78–86, December 2004, ISSN 1536-1284. Available online at: http://doc.utwente.nl/55651/1/01368900.pdf.
75. J. Yuan, Z. Li, W. Yu, and B. Li, A cross-layer optimization framework for multicast in multi-hop wireless networks wireless Internet, in *Proceedings of the 1st International Conference on Wireless Internet, WICON05*, pp. 47–54, July 10–14, 2005, Budapest, Hungary.
76. K. Seada, M. Zuniga, A. Helmy, and B. Krishnamachari, Energy-efficient forwarding strategies for geographic routing in lossy wireless sensor networks, in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, Sensys04*,

pp. 108–121, November 2004, Baltimore, MD.

77. M. Chiang, Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control, *IEEE Journal on Selected Areas in Communications (IEEE JSAC)*, 23(1), 104–116, January 2005.
78. R. Madan, S. Cui, S. Lall, and A. Goldsmith, Cross-layer design for lifetime maximization in interference-limited wireless sensor networks, in *Proceedings of IEEE INFOCOM'05*, vol. 3, pp. 1964–1975, March 13–17, 2005, Miami, FL.
79. S. Cui, R. Madan, A. Goldsmith, and S. Lall, Joint routing, MAC, and link layer optimization in sensor networks with energy constraints, in *Proceedings of the IEEE International Conference on Communications, ICC'05*, vol. 2, pp. 725–729, May 16–20, 2005, Seoul, Korea.
80. A. Bohn, State of the art on energy efficient and latency constrained networking protocols for wireless sensor networks, Technical report, IDE0749, Halmstad University, Halmstad, Sweden, June 2007.
81. M. Younis and T. Nadeem, Energy efficient MAC protocols of ad hoc networks, Book chapter in *Wireless Ad-Hoc and Sensor Networks*, Chapter 9, ed., A. Safwat, Pub: Kluwer Academic Publishers, Hingham, MA. Available online at: http://tmrnadeem.ifastnet.com//Web_Page/html_css/papers/energy_chapter.pdf.
82. H. Karl and A. Wiling, Protocols and architectures for wireless sensor networks, Book Published by: John Wiley & Sons, Ltd, April 2005, ISBN: 13-978-0-470-09510-2.
83. A. Chandrakasan, S. Sheng, and R. Brodersen, Low power CMOS digital design, *IEEE Journal of Solid State Circuits*, 27(4), 473–484, 1992.
84. X. Xia and Q. Liang, Latency and energy efficiency evaluation in wireless sensor networks, in *IEEE 62nd Conference on Vehicular Technology, VTC-2005-Fall*, pp. 1594–1598, September 25–28, 2005, Dallas, TX.
85. A. Wang et al., Energy-efficient modulation and MAC for asymmetric microsensor systems, in *Proceedings of ISLPED 2001*, Huntington Beach, CA, August 2001.
86. V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, Energy-aware wireless microsensor networks, *IEEE Signal Processing Magazine*, 19(2), 40–50, 2002.
87. S. S. Kulkarni, TDMA services for sensor networks, in *Proceedings of 24th International Conference on Distributed Computing Systems Workshops*, (ICDCS 2004 Workshops), pp. 604–609, March 23–24, 2004, Hachioji, Tokyo, Japan.
88. A. Chandra, V. Gummalla, and J. O. Limb, Wireless medium access control protocols, *IEEE Communication Surveys and Tutorials*, 3(2), 2–15, 2000.
89. S. Singh and C. S. Raghavendra, PAMAS: Power aware multi-access protocol with signaling for ad hoc networks, *ACM Computer Communications Review*, 28(3), 526, 1998.
90. I. Chatzigiannakis, A. Kinalis, and S. Nikoletseas, Wireless sensor networks protocols for efficient collision avoidance in multi-path data propagation, in *Proceedings of the 1st ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks*, pp. 8–16, October 4–4, 2004, Venezia, Italy.
91. P. Havinga and G. Smit, Energy-efficient TDMA medium access control protocol scheduling, in *Proceedings of the Asian International Mobile Computing Conference, AMOC*, 31 October–3 November 2000, Penang, Malaysia.

92. M. Adamou, I. Lee, and I. Shin, An energy efficient real-time medium access control protocol for wireless ad-hoc networks, in *WIP Session of IEEE Real-Time Systems Symposium, RTSS01*, London, U.K., December 2001.
93. M. Salajegheh, H. Soroush, and A. Kalis, HYMAC: Hybrid TDMA/FDMA medium access control protocol for wirless sensor networks, in *The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC07*, pp. 1–8, September 3–7, 2007, Athens, Greece.
94. M. Younis, K. Akkaya, M. Eltoweissy, A. Wadaa, On handling QoS traffic in wireless sensor networks, in *Proceedings of the 37th Annual Hawaii International Conference on Computer Sciences, (HICSS'04)*, January 5–8, 2004, Big Island, HI.
95. D. Chen and P. K. Varshney, QoS support in wireless sensor networks: A survey, in *Proceedings of the 2004 International Conference on Wireless Networks (ICWN 2004)*, pp. 227–233, June 21–24, 2004, Las Vegas, NV.
96. R. Iyer and L. Kleinrock, QoS control for sensor networks, in *Proceedings of the 38th Annual IEEE International Conference on Communications, ICC'03*, vol. 1, pp. 517–521, May 11–15, 2003, Anchorage, AK.
97. Y. Liu, I. Elhanany, and Q. Hairong, An energy-efficient QoS-aware media access control protocol for wireless sensor networks, in *Proceeding of the 2nd IEEE International Conference on Mobile Ad-hoc and Sensor Systems, MASS'05*, Poster Session, November 7–10, 2005, Washington, DC.
98. W. L. Lee, A. Datta, and R. Cardell-Oliver, QMAC: A quality of service oriented medium access control protocol for data gathering in wireless sensor networks, Tech. Rep. UWA-CSSE-05-005, The University of Western Australia, 2005.
99. Q. Zhao and L. Tong, QoS specific medium access control for wireless sensor networks with fading, ACSP technical report TR-06-03-01, School of Electrical and Computer Engineering, Cornell University, June 2003.
100. Y. Wu, S. Fahmy, and N. B. Shroff, Optimal sleep/wake scheduling for time synchronized sensor networks with QoS guarantees, in *Proceedings of 14th IEEE International Workshop on Quality of Service (IEEE-IWQoS'06)*, pp. 102–111, June 19–21, 2006, New Haven, CT.
101. J. Frolik, QoS control for random access wireless sensor networks, in *Proceedings of the 5th IEEE Wireless Communications and Networking Conference, WCNC'04*, vol. 3, pp. 1522–1527, March 21–24, 2004, Atlanta, GA.
102. B. Yahya and J. Ben-Othman, An energy efficient hybrid medium access control scheme for wireless sensor networks with quality of service guarantees, Accepted and will appear in the *GLOBECOM'08 Proceedings, Ad-hoc and Sensor Networks Symposium*, New-Orleans, LA, 29 November–4 December 2008.
103. V. Bharghavan et al., MACAW: A media access protocol for wireless LANS, in *Proceedings of the ACM SIGCOMM '94 Conference on Communications Architectures, Protocols and Applications*, 24(4), 1994, August 31–September 2, 1994, London, U.K.
104. See http://www.zigbee.org/; A brife slide set on ZigBee entitled ZigBee overview, can be found under http://www.zigbee.org/en/resources.
105. LAN/MAN Standards Committee of the IEEE Computer Sciety. IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specfic requirements—

Part 15.4: Wireless medium access (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (LT-WPANs), October 2003.

106. A. Jhumka and S. Kulkarni, On the design of mobility-tolerant TDMA-based media access control (MAC) protocol for mobile sensor networks, in *LNCS Book Series Distributed Computing and Internet Technology Book*, ISBN:978-3-540-77112-8, Springer, Berlin/Heidelberg, November 2007.
107. J. F. Shi, X. X. Zhong, and S. Chen, Study on communication mode of wireless sensor networks based on effective results, in *International Symposium on Instrumentation Science and Technology*, 8–12 August 2006, Harbin, China, *Journal of Physics*, Conference series 48, 1317–1321, IOP Publishing Ltd, 2006. Available online at: http://www.iop.org/EJ/toc/1742-6596/48/1.
108. D. Kotz, C. Newport, R. S. Gray, J. Liu, Y. Yuan, and C. Elliott, Experimental evaluation of wireless simulation assumptions, in *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM04)*, pp. 78–82, October 4–6, 2004, Venice, Italy.
109. M. J. Handy, M. Haase, and D. Timmermann, Low energy adaptive clustering hierarchy with deterministic cluster-head selection, in *Proceedings of the Fourth IEEE Conference on Mobile and Wireless Communications Networks*, pp. 368–372, Stockholm, September 2002.
110. M. Leopold, M. B. Dydensborg, and P. Bonnet, Bluetooth and sensor networks: A reality check, in *Proceedings of 1st ACM Conference on Embedded Networked Sensor Systems (SenSys 2003)*, pp. 103–113, New York, November 2003.

第10章　无线传感器网络的定位技术

在无线传感器网络设计中定位是一个基本问题，位置信息能够在许多无线传感器网络应用中使用，例如事件检测，目标跟踪，环境监控和网络调度。另一方面，位置信息同样能够在不同的网络协议中使用，从而提高传感器网络的执行性能，例如，路由包使用基于定位的路由，使用几何方法控制网络的拓扑和覆盖，或者取得了路由的较好的负载均衡。手工配置位置或者为每一个传感器提供一个 GPS 系统对于大型传感器网络来说是昂贵的，不可行的。因此，发展传感器节点通过和已知位置的节点交换信息来计算自己位置的定位方法是重要的。这一章节将给我们介绍无线传感器网络的定位方法，包括理论知识和研究的挑战，以及对需要测离的方法和无须测距的方法进行讨论和比较。因为有了特定的硬件，基于距离的模可以根据点到点的距离测量或者角度测量数据来获得更高的测量精度。另一方面，无须测距的机制可以通过花很少的花费获得粗糙的定位精度。在章节的最后，读者将会了解到无线传感器网络中最重要的设计和定位技术的最新研究趋势。

10.1　概述

在无线传感器网络中定位的最主要任务是获得每个节点绝对或相对，精确或近似的位置。下面获得的定位信息是有用的，甚至在无线传感器的许多方面都是必需的。

1）当特定的事件被传感器节点发现时，提供该事件的位置信息。

2）使用传感器网络提供一串位置序列用来跟踪一个正在移动的对象。

3）为每个单独的传感器获得的测量数据提供位置标记。

4）使用活跃节点的位置信息来确定它们覆盖质量。

5）基于几何技术控制网络拓扑来维持连通性以及节约能源。

6）根据当前节点，它的邻居节点和目的地的位置来制定路由决策，使用基于位置的路由选择来减少路由代价。

7）在路由协议中通过基于位置信息广播通信来获取负载均衡。

8）为多种多样的位置感知服务提供使用者的位置信息，例如找到附近的服务器或打印机。

9）支持 geocast[22] 和 mobicast[16] 服务，使得用户能够在特定的时间向特定的区域发送信息。

然而，在无线传感器网络中定位是一项极具挑战性的工作。手工配置位置和为每个传感器提供一个 GPS 系统是两种直观的方法。在小规模、固定的无线传感器

网络中，每个传感器节点的位置是固定的，因此它们能够被手工部署。然而当我们考虑大型的无线传感器网络时，手工配置位置是不可行的。卫星GPS是最有名的户外定位系统。然而，为每个微小的传感器配置一个GPS系统是昂贵的。而且，每个传感器只有有限的电池能量，不能够支持GPS系统的运行。最后，由于建筑物会阻挡卫星信号，GPS系统在室内是不适用的。因此，开发新的经济可行的无线传感器网络定位方法是非常必要的。

最近，传感器网络的定位问题已经引起了许多研究者的注意，而且提出了大量的定位方法。这些方法可以概括为两方面：基于测距的方法和无须测距的方法。基于测距的方法利用收到信号的强弱（Received Signal Strength，RSS）、到达时间（Time of Arrival，ToA）、时间差（Time Difference of Arrival，TDoA）、到达角度（Angle of Arrival，AoA）来测量节点间的距离和角度，然后使用这些距离和角度计算节点的位置。无须测距的方法不使用上面的测量技术，它们使用替代的方法，例如采用基于跳数的方法来定位节点。在这一章，我们将简略的回顾一些属于这两类的最近发展的定位方法。大型无线传感器网络的定位要实现以下的要求：定位方法必须是精确的（也就是说，估算位置和真正的位置之间的误差要小），分布式的（即不依赖于全局基础设施），健壮的（就是说能够容忍节点失效和测量错误）和高效的（就是具有较少的计算和通信代价）。然而，同时满足所有的要求是非常困难的。当前的方法通常只能满足其中一个或几个要求。

在10.2节介绍一些定位问题的理论基础。在10.3节和10.4节分别讨论目前基于测距和无须测距的方法。10.5节在比较所有方法之后给出一个总结，并指出在定位中一些尚未解决的问题。

10.2 理论基础

在详细讨论定位方法之前，我们首先介绍一些理论基础，比如基本的距离测量方法，三边测量，三角测量和关于定位的固定理论。

10.2.1 距离测量

为了确定一个传感器节点的位置，第一步是测量该节点和邻居节点（或参考节点）之间的距离或角度。为无线传感器网络提出各种各样的测量方法。

（1）到达时间（ToA） ToA方法通过测量信号在两个节点间的传输时间来估计两个邻近节点间的距离。如果信号的传播速度是已知的，距离等于时间乘速度。ToA方法通常采用传播速度低的信号，例如声波和超声波信号。无线电波可用于ToA，但是结果不是很精确。信号的传播速度也依赖于外部因素，例如温度、压力和湿度。这些可能会影响ToA测量的精度。

通常，ToA方法需要时间同步。为了避免时间同步，ToA需要测量乒乓球式的

往返行程时延。例如，我们想估计 a、b 两节点间的距离，信号的传播速度为 v。如图 10-1a 所示，在 t_a 时，a 发送一个信号到 b；b 在 t_b 接收到该信号，在 t'_b时回发信号，a 在 t'_a时收到回发的信号。然后可以用以下公式计算 a 到 b 的距离：

$$|ab| = \frac{((t'_a - t_a) - (t'_b - t_b))v}{2}$$

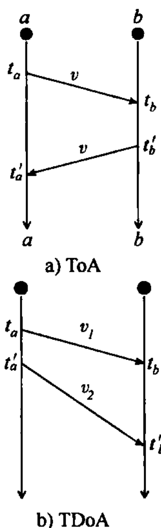


图 10-1　距离测量

（2）到达时间差（TDoA）　TDoA 方法和 ToA 方法相似，因为它们都是利用信号在两个节点间的传输时间。然而，不同于 ToA 方法，TDoA 采用两种传播速度不同的信号，例如，超声波和射频信号，来克服精确同步或者乒乓球式双向握手的需求。假使这两种信号的速度分别为 v_1 和 v_2。再一次，假设我们想估计 a、b 两点之间的距离。如图 10-1b 所示，a 在 t_a 和 t'_a分别发送两种信号；b 分别在 t_b 和 t'_b收到信号。我们知道 a、b 两点间的距离等于 $v_1(t_b - t_a)$ 和 $v_2(t'_b - t'_a)$。换句话说，$\frac{|ab|}{v_1} = t_b - t_a$ 并且 $\frac{|ab|}{v_2} = t'_b - t'_a$。为了避免同步，我们采用两个方程之差，$|ab|\left(\frac{1}{v_1} - \frac{1}{v_2}\right) = (t'_b - t_b) - (t'_a - t_a)$，因此可用下面的公式计算距离：

$$|ab| = ((t'_b - t_b) - (t'_a - t_a))\left(\frac{v_1 v_2}{v_1 - v_2}\right)$$

与 ToA 方法相比，TDoA 一般可以获得更好的准确度，但是需要每个节点至少有两种类型的信号发送器和接收器。

我们注意到，这里的 TDoA 的定义不同于发射器定位技术所使用的 TDoA[1]。在发射器定位问题上，TDoA 是加上频率差（Frequency Difference of Arrival，FDoA）来被动地估计固定发射器的位置。它包括两个分开的接收器的到达时间之间的测量差。然而，在这里，TDoA 是作为两种不同信号到达相同接收器的到达时间差的测量值。

（3）接收的信号强度（RSS）　当信号在介质中传播，它们的信号强度随着传播距离的增大而减小。根据这一事实，在接收器处接收到的信号强度可以用来计算接收者和发送者之间的距离。在理想的情况下，无线电传播路径损耗模型可以表示为下列公式：

$$P_r = c\frac{P_t}{d^\alpha}$$

式中　P_r——接收器接收的信号强度；

P_t——发送器发送的信号强度；

α——路径耗损系数，根据具体的环境通常是2~5的一个常数；

c——一个常数；

d——发送者和接收者之间的距离。

因此，给出接收信号的强度、P_t、α、c，我们可以用下面的公式计算出距离：

$$d=\sqrt[\alpha]{\frac{cP_t}{P_r}}$$

基于接收信号强度的方法的优点是传感器节点不需要附加的硬件和通信能耗。但是这种方法的准确度受到路径损耗模型准确度的限制。通常，使用上述简单的路径损耗模型可能会出现严重的错误，因为接收信号强度是非常复杂的，而且即使是固定的发送器和接收器，信号也是多样的和振荡的。这是由于受到各种不可控制因素的影响，例如，受到多径衰落、障碍物和地形的影响。因此，如果基于接收信号强度的方法要获得一定的准确度，需要使用更多复杂的技术。

（4）到达角度（AoA） 上面三种方法都是关注于直接测量距离，但是另外一种替代方法是测量与参考节点（位置已知的）之间的角度或者接收信号的方向，来替代距离测量，从而确定节点的位置。基于AoA的方法通常是使用定向天线或阵列天线来测量设备和参考节点连接线的角度。在二维空间，需要两个测量角度（两个参考节点的角度）和一个长度测量（例如两个参考节点之间的距离）来定位一个节点。

（5）到达频差（FDoA） FDoA是一种实现发射器精确定位的技术[1]。它涉及两个接收器从一个发送器接收到的接收频率之间的测量差异。因为接收频率的差异是多普勒频移的差异引起的，FDoA也称为微分多普勒。最近，有一种应用FDoA来定位传感器的趋势[3,19,20,23]。在参考文献［19］中，传感器通过测量从一个移动锚发出的一个音调的声波多普勒频移来估计自己的位置。它假定传感器知道自己的位置，锚的头以及声音音调的频率。在参考文献［23］中，两个锚，一个主机和一个辅助机，以略微不同的频率发送一个纯正弦波，频率的干扰导致低的包摆动频率，它的相位可以通过在特定时间内由两个接收器来测量。这个相位差是发送器和接收器之间距离的线性组合。在参考文献［3，20］中，使用一种相似的技术来跟踪移动的传感器。但不是测量相位，它测量多普勒频移，当传输信号源相对于观察者运动时进行测量。因为多普勒频移是由发送器和接收器的相对速度决定的，移动节点的绝对速度可以通过使用传输频率的先验知识和已知位置的静态传感器的观察频率来得到。

10.2.2 三边测量

三边测量是定位系统最基本的一项技术，并已使用多年。三边测量使用两个或多个参考节点的已知位置，而且节点之间的测量距离需要位于每个参考节点。为了

单独使用三边测量来准确，唯一的确定二维平面上一个节点的相对位置，一般需要三个参考节点。

三边测量的基本思想如下。为了定位坐标为（x_d，y_d）的节点 d，我们需要三个参考节点，它们的坐标分别为（x_a，y_a）、（x_b，y_b）、（x_c，y_c）。如图 10-2a 所示，d 点的位置应该设在 a、b、c 三个圆的交叉点上。因此有如下三个方程：

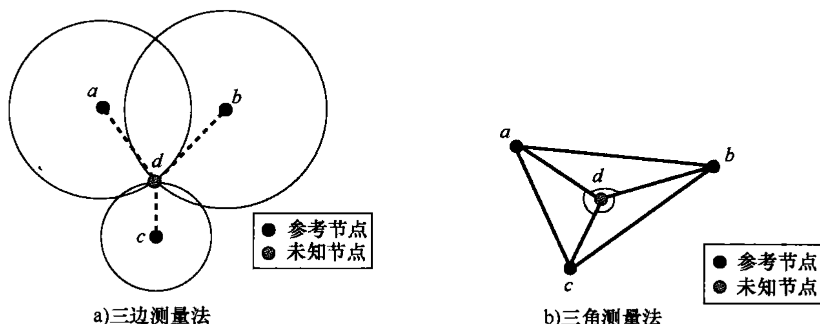


图 10-2　定位算法的基础

$$|ad|^2=(x_d-x_a)^2+(y_d+y_a)^2$$
$$|bd|^2=(x_d-x_b)^2+(y_d+y_b)^2$$
$$|cd|^2=(x_d-x_c)^2+(y_d+y_c)^2$$

通过求解上面的等式，可求得 d 的位置。三个参考节点 a、b、c 不能在一条直线上。

三边测量也可以应用于三维网络。唯一的区别是定位一个三维网络中的节点需要四个参考节点而不是三个。

10.2.3　三角测量

在三角法和几何法中，三角测量是求一个点的坐标的过程，它通过使用由这个点和两个或三个已知的参考节点构成的三角形的角和边的测量值，并利用正弦和余弦求得。不像三边测量只使用距离测量，三角测量还使用角度测量，如基于 AoA 的方法。例如，在二维网络中，如果知道一个节点和三个已知的参考节点之间的三个角度的话，我们可以非常容易地计算出这个节点的位置。

图 10-2b 显示了一个例子。我们再次假设有三个参考节点 a、b、c。未知节点 d 能测量∠adb、∠bdc 和∠adc 三个角度。因为三个参考节点的位置都已知，因此它们之间的距离也可求得。然后我们可以得到以下等式：

$$|ad|^2=(x_d-x_a)^2+(y_d+y_a)^2$$
$$|bd|^2=(x_d-x_b)^2+(y_d+y_b)^2$$
$$|cd|^2=(x_d-x_c)^2+(y_d+y_c)^2$$

$$|ab|^2 = |ad|^2 + |bd|^2 - 2|ad||bd|\cos\angle adb$$
$$|bc|^2 = |bd|^2 + |cd|^2 - 2|bd||cd|\cos\angle bdc$$
$$|ac|^2 = |ad|^2 + |cd|^2 - 2|ad||cd|\cos\angle adc$$

通过求解以上的6个等式，可求得 d 的坐标（x_d，y_d）。

10.2.4 网络定位理论：定位和固定理论

当有足够的测量数据可用时，只要通过使用参考节点测得的距离和角度，三边测量法和三角测量法都可以计算传感器节点的位置。然而，在实践中，由于无线传感器网络的测量限制，很多节点的位置都不能被唯一确定。最重要的问题是在什么条件下传感器网络的定位问题是可解决的（也就是说每个节点都有一个唯一的定位策略）。单一网络定位问题和称为固定理论的数学主题有着紧密的联系[15]。最近，在传感器领域公布了几个最新的研究成果[4,10,14]。

与距离信息有关的网络定位问题是：给定网络图 $G=(V, E)$，在 d 维空间 R^d 中的参考节点 p_j 的位置和每对邻近节点对$(i, j)\in E$之间的距离，在真实的 d 维空间 R^d 中确定所有节点 $v_i \in V$ 的位置 p_i。只要在 d 维空间 R^d 中恰好有一组未知节点，与它们有关距离和位置的有用信息是一致的，那么定位问题可以解决或者网络是可定位的。定位问题也可以描述成一个点的形成，$F_p=(\{p_1,p_2,\dots,p_n\},L)$，在这里 p_i 是节点 i 的位置，L 是一组具有节点间距离的链接集（包括未知节点与参考节点之间以及参考节点和参考节点之间的距离）。然后，在参考文献［4，10］中，下面的定理给出了定位网络的条件。

1. 定理10.1

对于一个位于 d 维空间 R^d 的网络，$d=1$，2，3。如果在一般情况下至少有 $d+1$ 个参考节点，只要对于图 G 每个节点的形成都是全局固定的，那么网络是可以唯一定位的。

下面我们给出一个全局固定的大致定义。随着点连接成边代表了距离的限制，我们来考虑一个点的形成（和其对应的图）。如果没有其他包含不同点和保留所有的距离限制的点形成，那么我们称这个点形成或与它对应的图是全局固定的。如图10-3所示，给出了在二维平面中非全局固定图和全局固定图。

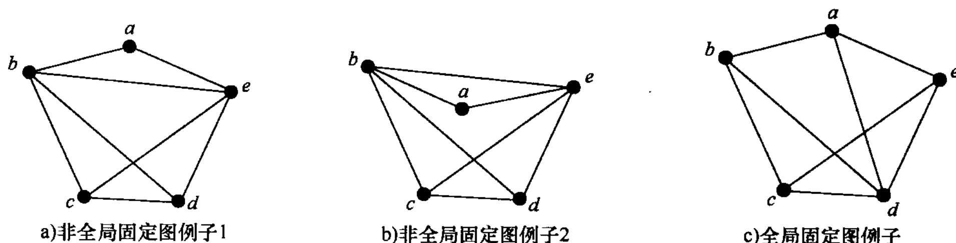


图10-3 两个非全局固定图和一个全局固定图的例子

固定理论给出了一个有效的方法来检查一幅图是否是全局固定的。下面的理论给出了在二维空间测试全局固定的必要条件。

2. 定理 10.2

在二维空间中一个节点数 $n \geqslant 4$ 的图是全局固定[17]，当且仅当它是固定冗余并且在二维空间 R^2 上是三相连的。

如果删除任意一条单独的边此图还是全局固定的，那么这图是固定冗余的。为了测试一般固定，可以使用下面的定理。

3. 定理 10.3

在二维空间 R^2 中，具有 n 个顶点和 $2n-3$ 条边的图是全局固定的，当且仅当没有超过 $2n'-3$ 条边的子图，n'为子图的顶点数。

注意，上面两个定理只适用于二维空间，不适用于更高维的空间。

尽管可以有效地确定全局固定，全局固定加权图的实现问题（网络定位问题）是 NP 难的。Aspneset 等人[4]证明了，只要给全局固定加权图提供分区问题的一个多项式时间归约就可以解决这个问题。有关定位问题的详细信息，请查看参考文献 [4]。

10.3 基于距离的定位方法

在本节，我们将开始探讨现有的几种无线传感器网络的定位方法。首先，我们关注基于距离的定位方法，这个方法用精确的距离和角度的测量技术 ToA、TDoA、AoA 和 RSS 估计节点之间以及节点和参考节点之间的距离。接着，我们称那些位置已知和可以向其他节点发送信标消息的参考节点为锚节点或信标节点。我们把基于距离定位的方法分为 4 类：单跳锚方法，多跳锚方法，移动锚使用方法和不使用锚的方法。

10.3.1 单跳锚方法

如果锚节点具有足够的能量，那么每个传感器节点都可以和它们通信，这种定位方法是线性的。如果可以在二维空间接收到至少 3 个锚节点发送的信息或者在三维空间可以接收到至少 4 个锚节点发送的信息，那么三边测量和三角测量都可用于传感器的定位。下面，我们只讨论一些用这种方法实现的例子。

（1）GPS　GPS 是最著名的使用单跳锚的户外定位系统。GPS 利用至少 24 个中轨道卫星发送精确的微波信号，使得 GPS 接收器能够确定自己的位置、速度、方向和时间。GPS 接收器可以使用 4 种 GPS 位置信号，通过简单的三边计算得到自己在三维空间中的位置。目前，GPS 定位系统的准确度大约为 3 - 20m。然而，GPS 对于无线传感器网络也有一些明显的缺点。为每一个节点配备 GPS 系统是昂贵的，而且 GPS 在很多情况下是不可用的，例如室内和水下环境。

（2）雷达系统　雷达系统[5]是利用在多基站的射频信号强度信息执行一个定位服务，这个多基站在某个影响区域提供重叠的有效范围。它使用射频信号强度（RF Signal Strengh，RSS）作为发送器和接收器之间距离的指示器。三边测量法使用这个距离信息来定位使用者。而且，雷达系统通过搜索来自离线阶段的录制场景来提高精度。在离线阶段，这个系统为已知的发射阵地在一套固定的接收器上建立了一个射频信号强度数据库。在正常运行阶段，一个发送器的射频信号强度被一个固定的接收器测量，并送到中央电脑，检查射频信号强度数据库来获得当前发送器的最合适位置。

（3）Cricket定位系统　Cricket[27]是为在室内，移动的，与位置有关的应用程序服务的定位系统。Cricket系统使用遍布在建筑物周围的多个锚节点，每个锚节点使用无线电和超声波信号产生信标。当每个节点收到多个锚节点发送的信号，它使用TDoA方法来估计自己到每个锚节点的距离，然后通过三边测量计算位置信息。在参考文献［27］中，作者也考虑来自不同锚节点的信号冲突和干扰，并描述了一种随机算法来克服这一影响。

（4）AoA法　Nasipuri和Li[25]提出了一种技术，这种技术中每个传感器节点通过获得和固定锚节点的相对角度来确定自己在传感器网络中的位置。它们认为每个锚节点具有向整个网络发送无线信标信号的特殊能力，而且每个信标信号由一个以恒定角速度旋转的窄波束上的连续射频波信号组成。然后每个节点可以记录收到不同锚信号时的时间，通过三角测量的方法估计它的角度和相对于锚节点的位置。在理想情况下，这种定位技术要正常工作至少需要有三个锚节点，额外的锚节点可以用来解决多路传输带来的错误。这种被提议的定位方法具有很好的准确度，而且在传感器节点处需要极少的额外的复杂性。主要的误差原因来自于方向信标信号波束的宽度。然而，已发现的定位误差在波束宽度在15°以内时很小。而且这种方法的效果不受网络中传感器节点密度的影响。

（5）水下传感器网络定位法　在水下环境[2]定位比在陆地上定位更困难。由于射频信号会被水吸收，所以不能使用射频信号。因此，一种声信号成为水下定位的选择。声信号具有以下的特点：低带宽，高时延和高误比特率。因为声信号的传播速度会随着盐度、压力、温度的变化而变化，所以在水下获得节点间的准确距离是困难的。我们提出了一些用于水下传感器定位的方法[8,9,34]。它们大多数都使用陆地上传感器定位的方法，不同的是两点：①使用声波；②需要更多的锚节点（因为在水下的传感器网络是三维网络）。

10.3.2　多跳锚方法

前一节我们讨论了如果一个传感器节点可以接收到来自单跳链接的信号，那么应如何定位一个传感器节点。然而在许多传感器网络中（尤其是大型的传感器网络），网络由多跳链接组成，许多节点都不能直接和锚节点通信。因此，多跳锚方

法必须被使用。

10.3.2.1　迭代和协作多点监视

在二维网络中，如果可以接收到至少三个节点发送的信号，那么三边测量可以定位一个传感器节点。然而，许多节点都不能直接和锚节点通信来计算它们的位置。多点监视的方法已被推荐用于解决这个问题。多点监视法的基本思想是节点测量它们到邻居节点的距离，并和邻居节点分享自己的位置信息来协作地计算它们的位置。在参考文献［29，30］中，Savvides 等人建议使用迭代的多点监视和协作的多点监视方法。

在多点监视的第一阶段，每个传感器节点通过使用 ToA、TDoA 或者 RSS 方法来测量和邻居节点之间的距离。如果有足够的邻居节点，而且它们的位置都是已知的，那么三边测量可以用来计算节点的位置。因此，在最后阶段，只有那些能够和足够的锚节点直接通信的节点才可以获得它们的位置信息。

在迭代的多点监视方法[29]中，传感器节点的位置在第一阶段已被唯一确定，并发送它们的位置信息给邻居节点。这些节点可看做是新的锚节点。如果一个节点不知道它的位置，但一旦有了足够多的知道位置的节点，它就可以获得它的位置，并把它的位置信息发送给所有的邻居节点。这个迭代的过程将持续到没有节点再可以进一步定位（也就是说，在二维空间中未知位置的节点没有三个已知位置的邻居节点）。迭代多点监视法的缺点是把定位好的节点作为锚节点，这会引进大量累积误差。

为了进一步解决在迭代结束时有些节点没有足够的锚节点来定位的问题，协作的多点监视法[29]被引入。它试图通过使用多跳的位置信息来估计位置。它决定了包含锚节点和未知位置节点的网络协作子图，以便它们的位置和节点间的距离可以通过某些优化算法解二次方程组得到。在参考文献［30］中，作者也介绍了使用最小二乘估计的方法来完善节点的位置。

10.3.2.2　扫描法

最近，Goldenberg 等人[13]引入了一种为稀疏网络定位的扫描技术。在实际的传感器网络中只有一些传感器节点可以被唯一确定，还有一些传感器的节点无法被唯一确定。然而，许多这样的节点可以定位到一组可能的位置中。在许多情况下，知道一组所有可能的位置对于传感器来说是很有用的。这种定位被称为有限的定位而不是唯一的定位。扫描定位的思想与迭代或协作多点监视是相似的。算法的每一阶段，不仅使得可以定位的节点被定位，而且每一个可以有限定位的节点都会产生一组可能的位置。这些信息可以在邻居节点间交换并用于去除一些可能的位置。这种扫描方法的一个缺点是可能的位置会随着节点数的增加而呈指数级的增加。图 10-4 显示了一个在扫描法下减少可能位置的例子。

假设节点 a、b 是锚节点。节点 w 已测量出到 a、b 的距离。因此节点 w 有两种可能的位置 w 和 w'。另一个节点 u 获得了到 w、b 的距离之后，u 有四种可能的位置：

u、u'，它们来源于w、b；u_1'、u_1，它们来源于w_1、b；然后，如果节点v知道它的位置并向邻居节点广播，那么u将收到v的位置信息，并能够移除不可能的位置（u'和u_1）。

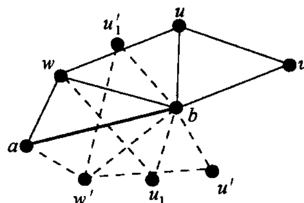


图 10-4 在淘汰法中消除可能位置的说明（来自 Goldenberg，D. K. 等人，稀疏网络中的定位使用淘汰，第12次移动计算和网络年度国际会议）

10.3.2.3 多维排列

多维排列（Multidimensional Scaling，MDS）是用于分析一组对象上不同数据的技术。最近它已经被用于在几何空间根据它们的成对距离恢复邻居传感器节点的位置[18,31]。基本上，MDS比较锚节点的估计位置和它们的真实位置，并通过迭代调整优化所有传感器节点的估计位置。MDS法也用于分布式的网络。这种方法见参考文献［31］，包括以下几个阶段。

在第一阶段，为每个节点形成邻近传感器节点的局部地图。局部地图是一幅包括所有到自身的距离为常数跳的传感器节点的地图。根据局部地图，可以计算每对节点间的最短距离。这些距离为多维排列形成距离矩阵。这一步的时间复杂度是$O(k^3)$，k是图中的平均节点数。多维排列使用距离矩阵来获得最大的特征值和特征向量，它们可以用来建立一幅局部地图。而且，这一步的时间复杂度也是$O(k^3)$。

在第二阶段，多个局部地图合并成一幅整体分布图。有很多方法可以合并当地分布图。例如，局部图可以任意地选为核心图。然后核心图加上邻居节点的局部图组成一幅新的核心图。这一过程将持续到核心图包括整个网络。这一步的时间复杂度是$O(k^3 n)$，n是传感器节点的个数。之后在这阶段的最后，我们获得所有节点的位置当然也可能包含错误。

在第三阶段，估计的位置可以通过最小二乘法优化。在最后阶段，如果有足够多的锚节点，那么所有节点的位置可转化为绝对位置。

10.3.3 移动锚应用法

在单跳锚和多跳锚方法中，因为没有足够多的锚节点，一些传感器节点的位置无法被唯一确定。因此一些定位方法部署了大量的锚节点来避免这一情况。然而，这带来了更多的开销。一种有效解决这个问题的方法是引入移动的锚节点（也叫做移动信标）。这些移动信标可以在网络周围移动，并在特定或任意时刻知道它们的位置（不是通过GPS或预定义的位置）。移动锚发出的信标信号包含它的位置信息。当一个节点接收到来自移动锚的信标信号，更多用于定位的信息将被获得，因此可以提高定位的准确度。不同的使用移动锚的方法被提出[11,12,32,38]。在这里我们简单地回顾一下水下传感器网络定位的方法。在参考文献［12］中，一种水下自主航行器（Autonomous Underwater Vehicle，AUV）被用于帮助水下传感器定位。当水下自主航行器在水面时，它可以收到来自GPS的信号，并能够获得它的位置。然后它潜入水下并遵

循已知的轨迹航行。当在节点间移动，它广播了一些信息，其中包含它自己目前的情况。这些信息可以作为信标信号。当一个节点收到来自水下自主航行器的信号时，它可以测量它到水下自主航行器的距离，并获得水下自主航行器的位置，如果这个节点获得足够多的这样的信息，它就可以通过三边测量计算它自己的位置。水下自主航行器通常遵循晶格轨迹、螺旋轨迹或其他轨迹。然而，决定它使用哪种轨迹是困难的，因为不知道未知节点的位置。一种简单的方法（潜水“N”上升，移动信标在水中下沉和上升并广播它们的位置）在参考文献［11］中被提到。

而且基于多普勒频差的方法[3,19,20]同样使用移动信标。

10.3.4 无锚节点法

到目前为止我们所讨论的所有定位方法，通过 GPS 来定位节点就是假设锚节点或移动信标来定位节点。还有另外一套方法[6,24]可以不使用任何锚节点，唯一的目标是获得相对位置而不是绝对位置。不使用锚节点的定位方法（例如参考文献［24］）通常包括下面三个主要阶段。在第一阶段，建立局部的坐标系统。每一个传感器节点是一个簇的中心，它测量其到相邻节点的距离。相邻节点同样测量自身到其他节点的距离，并向相邻节点发送消息，其中包括它们的位置消息，并发送包括它们到相邻节点距离的消息。第二阶段在每个簇中优化传感器节点估计的位置。不同的最优化技术，如弹簧发送或牛顿迭代，随着测量距离的限制可以被使用。在第三阶段，通过找到两个簇之间的共同节点集和解决旋转，反射，可能的最好对齐簇，相邻节点的局部坐标系统的变化可以被计算。不使用锚节点法的优点是每个节点都有一个以自身为原点的本地坐标系统。而且，传感器节点只可以测量它到邻近节点的距离并向邻近节点发送信息，因此这些方法是完全分布式的。

10.4 无须测距的定位方法

上面所提到的基于测距的方法，都能够为传感器节点提供一个精确的位置。然而，在一些应用领域，由于传感器节点上的硬件成本和限制，通常不使用基于测距的定位方法，这些方法是依靠点到点的绝对距离估算。因为粗准确度足以满足大多数传感器网络应用，无须测距的方法以经济高效替代基于测距法的昂贵的代价。无须测距法被分为两类：基于跳数的方法和基于区域的方法。

10.4.1 基于跳数的方法

在基于跳数的方法中，锚节点通常是放在边界或只在一个区域的角落。

10.4.1.1 基于距离向量的定位

与距离向量路由相似，在基于距离向量定位的方法中，每个传感器（节点）只能和它邻近的传感器通信，并且每个传感器（节点）估计它到锚节点的距离，并向

它的邻居节点发送信息。

最基本的基于距离向量的定位方法是距离向量跳广播的方法[26]，它由三个阶段组成。第一阶段，采用经典的距离向量交换，因此每个传感器节点可以计算它到锚节点的距离间的跳数。第二阶段，每个锚节点可以计算一跳的平均距离。记住在第一步之后，每个锚节点可以得到到任意锚节点的跳数距离。因为锚节点的位置是事先知道的，每一个锚节点可以计算每一跳的平均距离。然后锚节点通过距离向量交换广播信息。最后，在收到一跳的距离后，传感器节点可以估算它到锚节点的距离。然后传感器节点可以通过三边测量获得它们的位置。当一个锚节点j计算一跳的平均距离时，它遵循下面的公式：

$$c_i = \frac{\sum_j \sqrt{(x_i - x_j)^2 + (x_i - x_j)^2}}{\sum_j h_{ij}}，\text{对于所有其他的锚节点}，j。$$

式中　c_i——到锚节点i一跳的平均大小；

(x_i, y_i)和(x_j, y_j)——锚节点i，j的坐标；

h_{ij}——i，j间的跳数。

如图10-5所示，图中有三个锚节点a、b、c，因为这些节点的位置都事先知道，所以，它们知道它们之间以米为单位的距离。在距离向量交换之后，a可以知道它到b和c的跳数分别为2和6。因此根据上面的公式，a可以计算一跳的平均距离，就是$(100+40)/(6+2)=17.5$。同样的，锚节点b和c也可以计算它们一跳的平均距离，分别为16.42和15.90。锚节点通过距离向量的交换广播它们的大小。传感器节点可以利用这些值计算它们到每一个锚节点的距离。基于距离向量的方法的优点的使用非常简单，但是它只能在各个方向上的性质都相同的各向同性的网络中工作。

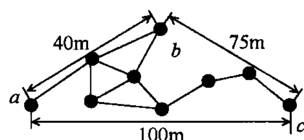


图10-5　DV一跳：估计一跳长度（来自Niculescu，D. 和Nath，B.，J. Telecommun. Syst，2003，22（267）.

参考文献［26］提出除了一跳的方法，还有在基于距离向量的方法中还有另外三种多跳的距离广播方法。第二种方法是基于距离向量和距离的广播方法，它和基于距离向量和跳数的广播方法很相似。这两种方法的唯一的区别就是在距离向量交换中，前者是使用测量的距离而不是跳数。这里，节点间的距离是通过ToA、TDoA、RSS或者AoA计算得到的。基于距离向量和距离的广播方法比基于距离向量和跳数的广播方法更精确。但是这种方法实际上不是无须测距的方法。在第三种方法中，欧几里得广播方法代替了距离测量，实际的到锚节点的欧几里得距离被广播。这个实际的距离可以从GPS系统获得。最后的广播方法是基于距离向量和坐标的广播方法，这里传感器节点的坐标被广播。这个方法和协作的多点监视相似。在广播坐标之前，每个节点应该建立它的局部坐标系统。当收到来自邻近节点的坐标，每个节点的坐标应该转变到自己的局部坐标系统中。

10.4.1.2　其他改进

通过基于跳数的方法获得的位置是比较粗糙的。根据基于距离向量的定位方法，更多的改进方法被提出来[28,33,36,37]。这里，我们简单的讨论一些。

在参考文献［28］中提出了一个两阶段的定位算法。第一个阶段主要是基于距离向量和跳跃数的方法。之后每一个节点获得它的估计位置，一个优化算法迭代的运行。每一个节点广播它们的估计位置，并收到邻近节点的位置和相应的估计距离。每个节点可以通过计算一个三角最小二乘解获得它的新位置。由于到邻近节点的距离约束将强迫新位置变为实际的位置。在几个迭代之后，位置更新变小，算法结束。

对于许多基于跳数的定位方法，估计位置的错误来自于节点到锚节点的距离。一般，锚节点部署在网络的边界，许多研究人员发现如果网络是均匀分布的，那么在网络的中心定位的准确度是最高的。基于这一发现，提出了选择性迭代多点监视的方法(SIM)[33]。SIM 选择具有较高的定位准确度的节点作为锚节点。基本上，SIM 根据所有锚节点的平均位置首先定位网络的中心。然后，围绕中心的节点是扮演新的锚节点的最好候选节点。模拟显示这种方法可以提高基于距离向量和跳数定位方法的准确度。

还需记住基于距离向量和跳数定位方法，在各向同性的网络中正常工作。然而，当节点分布不均匀时，节点之间的一跳的实际大小会变化很大。这会导致随着路径长度的增加而增大估计距离产生测量误差。在参考文献［37］中，Wong 等人提出密度感知的跳数测量方法（Density-aware Hop-count Localization，DHL)。DHL 考虑两个潜在的问题：在估计距离的过程中的密度和路径长度。在基于距离向量和跳数的方法中，一个节点到锚节点的距离是通过跳数乘每一跳的平均距离得到的。然而，在 DHL 中，每一跳的距离是基于局部的密度来计算的。由于距离估计的累积误差，估计距离由较少的跳数计算得到较为精确，因此可以得到一个较高的准确度等级。因此，DHL 到最近的锚节点的距离来估计位置会有较高的准确度。与这一想法相似，在参考文献［36］中，Wang 和 Xiao 提出了一种当在网络中有多个距离测量结果可用时，检查不准确的距离测量的方法。

10.4.2　基于区域的方法

在许多实际的应用领域中，我们没有必要得到传感器节点的坐标，我们只需要确定一个传感器节点分布的区域。因此基于区域的定位方法关注提供传感器节点分布区域的信息。

区域定位方法（Area Localization Scheme，ALS)[7]是一种集中的无须测距的方法，它提供传感器节点分布的区域。在 ALS 方法中，存在三种节点：锚节点、传感器节点和 sink 节点。ALS 通常将网络划分为格，并将锚节点分布在格的边沿。每一个锚节点采用不同的功率周期性地广播信标信号。因为以不同功率发送的信标信号可以到达不同的距离，也就是说，以高功率发送的信标信号可以到达更远的距离，锚节

点可以通过计算确定要到达不同距离所需的功率。在ALS中，每个锚节点设计一组不同功率的级别和可以覆盖整个网络的最大功率的信号。当每个锚节点广播信号时，这个信号包括锚节点的ID号和发送信号的功率。每个传感器节点只需监听和记录从锚节点接收到信号的功率。当接收到一个信号，每个传感器节点获得发送信号的锚节点的ID号和信号的功率级别。并且，传感器节点记录从每个锚节点收到的信号的功率级别和这个锚节点的ID号，然后发送信息给sink节点。sink节点将决定传感器节点分布的区域。

图10-6显示了一个例子，四个锚节点分别在正方形的四个角上。每一个锚节点会以三种不同功率级别广播信号。最低的功率级别记为整数1，最高的功率级别被记为整数3，级别为3时可以到达整个网络。〈3，1，3，3〉，〈1，3，3，3〉等是信号坐标。一个信号坐标用〈S_1，S_2，...，S_n〉表示，S_i是从锚节点i收到的最低信号级别。在图10-6中，轮廓线是每个功率级别都能到达的最远距离。很明显，每个节点的轮廓线可以被分为三个部分，所有的轮廓线把网络分成9部分。每一个区域都与一个不同的信号坐标相联系。很明显，在区域A传感器节点收到来自锚节点1、2、3的最低功率级别都是3，收到来自锚节点4的最低功率级别是2，因此这个区域A的坐标信号是〈3，3，3，2〉。

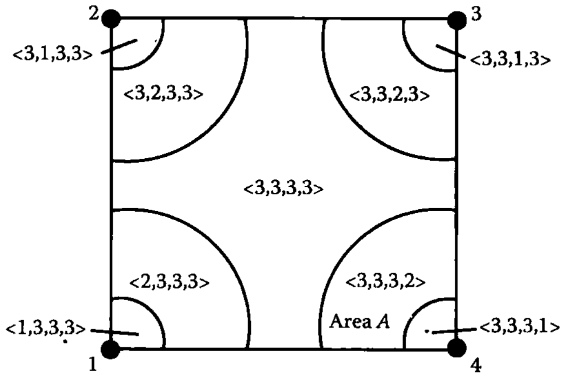


图10-6 ALS说明（获许使用。来自Chardrasekhar V and seoh W，对于水下传感器网络的区域定位模式。OCEANS 2006-ASia Pacific，2007）

Sink节点能够收到来自传感器节点的信号坐标。根据这个信息，Sink节点可以估计一个传感器的分布区域，并发送结果给传感器节点。ALS假设Sink节点事先知道锚节点的位置和它们的功率级别。Sink节点具有较高的计算能力。

在参考文献［35］中提出一种无须测距的定位算法，它被称为城市行人定位算法（UPL），是根据给出每个移动节点在市区出现的区域而提出的。UPL考虑两件事情：关于信息的障碍，例如墙和移动的节点。下面是UPL的基本思想。UPL假设移动的节点不能频繁收到锚节点发送的信息，移动的节点在市区是完全连接的。每一个移动的节点记录它分布的区域，并且当收到其他锚节点发送的信息时可以获得其他的分布区域。根据传输的距离，我们可以通过集合两个或更多的区域来计算节点的分布。由于每个节点的移动性，随着时间的推移，我们可以根据节点的速度和障碍物信息计算它存在的区域。

10.5　总结

定位用于确定所有传感器的位置，在传感器网络中是一个必要和基础的问题。在这一章，我们介绍了一些基本的定位技术，例如距离测量，三边测量和三角测量。并回顾从基于测距的方法到无须测距的方法的一些典型的定位方法。这些方法在精确或粗糙，分布或集中，是否放置锚等方面是完全不同的。表 10-1 总结并对它们作了比较。

表 10-1　定位方法比较

方法	准确度	算法	放置锚的位置	锚的百分比
全球定位系统	准确	分布式	人造卫星上	低（24 个）
单跳锚法	准确	分布式	边界	低或中等
多跳锚法	准确	分布式	随机或边界	中等
移动锚使用法	粗糙	分布式	移动物体	很少（1 或一些）
基于跳跃数法	粗糙	分布式	边界	低
基于区域法	粗糙	集中式	边界或网格	高

基于测距的定位方法通常会给出一个准确的位置信息，而无须测距的方法通常给出一个粗糙的位置信息。由于传感器网络分布式的特点，除了基于区域的定位方法，许多定位方法都是分布式的。单跳锚方法需要几个具有强大能力的锚节点能够覆盖更多的区域甚至是整个区域，或者有足够多的锚节点以至于每个传感器都能和它们中的一些通信。为了减少锚节点的数量或避免使用能力强大的锚节点，引入了多跳锚节点法。在这种方法中，锚节点所占的百分比是非常低的，锚节点通常是随机分布的。如果移动的节点被使用，锚节点的数量可以进一步减少甚至只有一个。对于基于跳数的方法，锚节点通常分布在网络的边沿，锚节点所占的百分比也很低。对于基于区域的方法，需要大量的锚节点用于覆盖整个网络，获得一定的准确度。

另外一个定位算法的重要标准是消息的开销，也就是说，在传感器节点间或锚节点间消息交换的数量。很明显，在 GPS 和单跳锚方法中消息的数量是最少的，因为每个节点只需监听来自锚节点发送的消息，而不用转发任何消息。唯一的消息是从锚节点广播的信标信号。然而，在多跳锚或基于跳数的方法中，特定的消息（例如位置、估计的距离或跳数）需要被交换，并且在邻近节点间广播。这必定会带来大量的消息开销。对于移动锚或基于区域的方法，消息的开销取决于系统的密度大小以及系统设计是如何实现的（例如传感器是否需要向其他节点广播消息或 sink 节点是否需要向接收端发送消息）。

尽管提出和研究了许多定位方法，但为了使分布式的传感器网络能够在恶劣的环境中拥有健壮和准确的算法仍然是一项艰巨的任务。在定位领域仍有许多问题需要解

决。例如，如何在测量出现错误的情况下保持定位的准确度。如何确保准确度、锚节点数量、时间、消息或能量开销之间的平衡。在三维或更高维的空间是否有足够的和必要的固定测试条件。最后，定位算法的设计严重依赖于传感器网络的应用。对于不同的应用的种种限制，需要精心设计合适的算法。

参考文献

1. David Adamy. *EW 102: A Second Course in Electronic Warfare.* Artech House, Boston, 2004.
2. I.F. Akyildiz, D. Pompili, and T. Melodia. Underwater acoustic sensor networks: Research challenges. *Ad Hoc Networks*, 3(3): 257–279, 2005.
3. I. Amundson, X. Koutsoukos, and J. Sallai. Mobile sensor localization and navigation using RF doppler shifts. In *Proceedings of 1st ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments (MELT'08)*, San Francisco, CA, 2008.
4. J. Aspnes, T. Eren, D.K. Goldenberg, A.S. Morse, W Whiteley, Y.R. Yang, B.D.O. Anderson, and P.N. Belhumeur. A theory of network localization. *IEEE Transaction on Mobile Computing*, 5(12): 1–15, 2006.
5. P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *Proc. of 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*, Tel Aviv, Israel, 2000.
6. S. Capkun, M. Hamdi, and J.-P. Hubaux. GPS-free positioning in mobile ad-hoc networks. In *Proc. of the 34th Annual Hawaii International Conference on System Sciences (HICCS)*, Maui, HI, 2001.
7. V. Chandrasekhar and W. Seah. An area localization scheme for underwater sensor networks. In *Proc. of the IEEE OCEANS Asia Pacific Conference*, Singapore, 2006.
8. V. Chandrasekhar, W.K.G. Seah, Y.S. Choo, and H. Voon Ee. Localization in underwater sensor networks: Survey and challenges. In *WUWNet '06: Proceedings of the 1st ACM International Workshop on Underwater Networks*, pp. 33–40, New York, 2006. ACM.
9. X. Cheng, H. Shu, and Q. Liang. A range-difference based self-positioning scheme for underwater acoustic sensor networks. In *Proc. of International Conference on Wireless Algorithms, Systems and Applications (WASA 2007)*, Chicago, IL, 2007.
10. T. Eren, D.K. Goldenberg, W Whiteley, Y.R. Yang, A.S. Morse, B.D.O. Anderson, and P.N. Belhumeur. Rigidity, computation, and randomization in network localization. In *Proc. of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, Hong Kong, China, 2004.
11. M. Erol, L.F.M. Vieira, and M. Gerla. Localization with dive'n'rise (DNR) beacons for underwater acoustic sensor networks. In *WUWNet '07: Proceedings of the Second Workshop on Underwater Networks*, pp. 97–100, New York, 2007. ACM.
12. M. Erol, L.F.M. Vieira, and M. Gerla. AUV-aided localization for underwater sensor networks. In *Proc. of International Conference on Wireless Algorithms, Systems and Applications (WASA 2007)*, Chicago, IL, 2007.
13. D.K. Goldenberg, P. Bihler, M. Cao, J. Fang, B.D.O. Anderson, A. Stephen Morse, and Y. Richard Yang. Localization in sparse networks using sweeps. In *MobiCom*

'06: *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, pp. 110–121, New York, 2006. ACM.

14. D.K. Goldenberg, A. Krishnamurthy, W.C. Maness, Y.R. Yang, A.S. Morse, and A. Savvides. Network localization in partially localizable networks. In *Proc. of 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005)*, Miami, FL, 2005.
15. J. Graver, B. Servatius, and H. Servatius. *Combinatorial Rigidity*. Graduate Studies in Math., AMS, 1993.
16. Q. Huang, C. Lu, and G.-C. Roman. Design and analysis of spatiotemporal multicast protocols for wireless sensor networks. *Telecommunication Systems*, 26(2–4): 129–160, 2004.
17. B. Jackson and T. Jordan. Connected rigidity martoids and unique realizations of graphs. *Journal of Combinatorial Theory, Series B*, 94(1): 1–29, 2005.
18. X. Ji and H. Zha. Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling. In *Proc. of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, Hong Kong, China, 2004.
19. R.J. Kozick and B.M. Sadler. Sensor localization using acoustic doppler shift with a mobile access point. In *Proceedings of IEEE/SP 13th Workshop on Statistical Signal Processing*, Bordeaux, France, 2005.
20. B. Kusy, A. Ledeczi, and X. Koutsoukos. Tracking mobile nodes using RF Doppler shifts. In *Proceedings of the ACM 5th International Conference on Embedded Networked Sensor Systems*, Sydney, Australia, 2007.
21. G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4(4): 331–340, 1970.
22. C. Maihofer. A survey of geocast routing protocols. *IEEE Communications Surveys and Tutorials*, 6(2): 32–42, 2004.
23. M. Maroti, P. Volgyesi, S. Dora, B. Kusy, A. Nadas, A. Ledeczi, G. Balogh, and K. Molnar. Radio interferometric geolocation. In *Proceedings of the ACM 3rd International Conference on Embedded Networked Sensor Systems*, Boulder, CO, 2005.
24. D. Moore, J. Leonard, D. Rus, and S. Teller. Robust distributed network localization with noisy range measurements. In *SenSys '04: Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 50–61, New York, 2004. ACM.
25. A. Nasipuri and K. Li. A directionality based location discovery scheme for wireless sensor networks. In *WSNA '02: Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 105–111, New York, 2002. ACM.
26. D. Niculescu and B. Nath. DV based positioning in ad hoc networks. *Journal of Telecommunication Systems*, 22(1–4): 267–280, 2003.
27. N.B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In *MobiCom '00: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, pp. 32–43, New York, 2000. ACM.
28. C. Savarese, J.M. Rabaey, and K. Langendoen. Robust positioning algorithms for distributed ad-hoc wireless sensor networks. In *ATEC '02: Proceedings of the General Track of the Annual Conference on USENIX Annual Technical Conference*, pp. 317–327, Berkeley, CA, 2002. USENIX Association.

29. A. Savvides, C.-C. Han, and M.B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *MobiCom '01: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, pp. 166–179, New York, 2001. ACM.
30. A. Savvides, H. Park, and M.B. Srivastava. The bits and flops of the *n*-hop multilateration primitive for node localization problems. In *WSNA '02: Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 112–121, New York, 2002. ACM.
31. Y. Shang and W. Ruml. Improved mds-based localization. In *Proc. of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, Hong Kong, China, 2004.
32. K.-F. Ssu, C.-H. Ou, and H. Christine Jiau. Localization with mobile anchor points in wireless sensor networks. *IEEE Transctions on Vehicular Technology*, 54(3): 1187–1198, 2005.
33. J.H.S. Tay, V.R. Chandrasekhar, and W.K.G. Seah. Selective iterative multilateration for hop count-based localization in wireless sensor networks. In *Proc. of 7th International Conference on Mobile Data Management (MDM 2006)*, Nara, Japan, 2006.
34. C. Tian, W. Liu, J. Jin, Y. Wang, and Y. Mo. Localization and synchronization for 3D underwater acoustic sensor networks. In *Proc. of 4th International Conference on Ubiquitous Intelligence and Computing (UIC 2007)*, Hong Kong, China, 2007.
35. A. Uchiyama, S. Fujii, K. Maeda, T. Umedu, H. Yamaguchi, and T. Higashino. Ad-hoc localization in urban district. In *Proc. of 26th IEEE International Conference on Computer Communications (INFOCOM 2007)*, Anchorage, AK, 2007.
36. C. Wang and L. Xiao. Locating sensors in concave areas. In *Proc. of 25th IEEE International Conference on Computer Communications (INFOCOM 2006)*, Barcelona, Spain, 2006.
37. S.Y. Wong, J.G. Lim, S.V. Rao, and W.K.G. Seah. Multihop localization with density and path length awareness in non-uniform wireless sensor networks. In *Proc. of IEEE 61st Vehicular Technology Conference (VTC 2005-Spring)*, Stockholm, Sweden, 2005.
38. C.-H. Wu, W. Sheng, and Y. Zhang. Mobile sensor networks self localization based on multi-dimensional scaling. In *Proc. of 2007 IEEE International Conference on Robotics and Automation*, Roma, Italy, 2007.

第 11 章　无线传感器网络中的数据聚合技术

无线传感器网络通常包括大量低成本的传感器节点，这些传感器节点严格限制了感知、计算和通信的能力。因为传感器节点的资源有限，所以最小化网络间的数据传输来提高传感器的平均寿命和整体带宽利用率是很重要的。在无线传感器网络中，数据聚合是传感器数据汇总和结合的过程，数据聚合是为了减少数据的传输同时增大数据的可靠性。本章调查了在无线传感器网络中当前最先进的数据聚合技术。在本章的第一部分，根据网络的拓扑结构，数据聚合协议被分为两部分。然后，因为数据聚合和安全协议都是无线传感器网络中必不可少的，因此正在研究它们之间的相互作用。在本章的最后，提出了在数据聚合问题方面的开放的研究领域和未来的研究方向。

11.1　概述

无线传感器网络由成千上万廉价的，低功率的，具有有限的计算和通信资源的感知装置组成，现在其越来越受到人们的欢迎[1]。这些网络为军事和民用的一系列问题提供高效的低成本的解决方案，包括战争环境监视、目标追踪、环境和健康的监测、野外侦查和交通管理。无线传感器网络技术的最直接应用就是通过低频率收集数据来监控远程环境。例如，边境地区可以通过使用上千个传感器轻松得以监控。在这个例子中，传感器节点将自动的形成无线互联网络并立即报告任何非法越境。不像传统的有线网络，部署无线传感器网络的费用是很低的。为了实现这一低成本的部署要求，传感器只有简单的硬件，在电池能量、计算能力、内存和存储方面都有限制。因为这些限制，要为数据聚合问题提供一个高效的解决方案是很困难的。在这些限制当中，电池能量限制是设计无限传感器网络协议时需要考虑的最关键的限制因素。为了解决传感器节点平均能量的消耗，提出了几个机制，例如，无线电调度、控制数据包消除、拓扑控制算法和数据聚合[2]。在这一章，我们关注于数据聚合技术，它的目的在于通过查看数据包的内容来结合和汇总来自一些传感器节点的数据包，从而实现能源效率。一个数据聚合方案的例子如图 11-1 所示，一组传感器节点报告来自目标区域的温度测量值。当基站查询这个网络来获得目标区域的平均温度时，它不是传输每个节点数据到基站，而是一个称为数据聚合器的节点，收集来自节点的温度读数并汇总这些数据（也就是计算均值），再发送计算后的平均温度值到基站。这个例子说明了这一点：数据聚合减少数据传输的数量，从而提高了在网络中的带宽和能源利用率。

在无线传感器网络中，如果当数据正在被发送到基站时，中间的传感器节点能够逐步地聚合数据，那么数据聚合的效益将增大。然而，当这些连续的数据聚合操作提

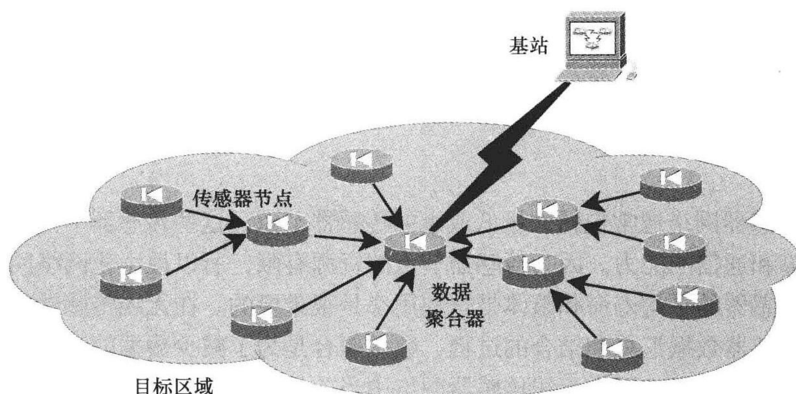


图 11-1 无线传感器网络中的数据聚合

高带宽和能源利用率的同时，它可能会对其他的性能指标产生负面的影响，例如时延、准确度、容错能力和安全性[2]。在这些指标当中，安全性是最重要的一个，因为大多数的无线传感器网络应用都需要一定程度的网络安全。而且，在安全性和数据聚合技术之间存在很大的冲突。安全协议要求节点加密，在数据传输之前要验证任何传输的数据，数据要被基站解密[28,30]。另一方面，数据传输协议往往会在每一个中间节点实现数据聚合，从而使能源效率最大化。由于这两个相互冲突的目标，在系统设计允许数据聚合但不牺牲安全性时，数据聚合和安全协议必须同时设计。实现数据聚合和安全的必要性使得很多研究者研究安全的数据聚合，因此这一点值得特别关注。

在这一章，我们给出在无线传感器网络中数据聚合模式的全面介绍。这一章的组织结构如下。首先我们从一个关于无线传感器网络的简单概述开始，给出数据聚合模式背后的动机。然后，根据网络拓扑，在两小节提出了当前最新的数据聚合研究。接着，我们研究数据聚合和安全协议之间的相互作用，因为它们对于无线传感器网络都是必不可少的。安全的数据聚合协议在两个部分被提出。首先，我们提出了一个解决方案，在这个方案中，在数据聚合之前传感器数据被加密，而且使用特定的机制来保证普通数据聚合的安全性。然后，在第二部分，安全的数据聚合协议能够通过对加密数据的解密进行数据聚合。在最后，我们提出了在数据聚合问题方面的开放性研究领域和未来的研究方向。

11.2 无线传感器网络概述

一个无线传感器网络是由分散在一定地理区域的大量传感器节点组成的。每个传感器节点的通信能力有限，信号处理和对网络数据的计算能力低。传感器节点由小型的电池供电，它们被认为是不可更换的。因此，传感器节点在能量来源方面也受到限制。上述传感器节点的资源限制和由大型无线传感网络引入的一系列新的研究问题和

挑战，这些问题和挑战是以前的研究不需要解决的。这些研究问题之一是旨在减少数据传输量的数据聚合。在给出数据聚合模式的详细介绍之前，我们在下面部分总结了无线传感器网络独特的特点，并提出了它们与数据聚合协议之间的关系。

（1）大规模　无线传感器网络的典型应用领域（如战场、栖息地的监测），需要一个很大的地理覆盖范围。同时，高密度的节点被需要用来对抗传感器节点的高的失效率、个别传感器节点读数的不精确和传感器节点有限的通信范围。某些应用程序需要 k 覆盖的方案，在这个方案中事件必须被至少 k 个传感器节点监测到。由于这些原因，无线传感器网络预计将扩大到上千个节点。来自这些大量的传感器节点的数据必须在传输到基站之前汇总，从而使由于数据传输而产生的能量消耗减少。数据聚合技术是一种汇总收集到的信息，而无须所有数据块的高效方法。

（2）资源限制　因为无限传感器网络低成本的实现要求，传感器节点只有一个简单的硬件，这就限制了传感器网络的计算和通信能力。例如，一个常见的传感器（*TelosB*）只有一个 16bit、8MHz 的 RISC CPU，这种 CPU 只有 10KB 的内存，48KB 的程序存储和 1024KB 的闪存[3]。而且一旦网络部署好了，传感器节点的电池更换或充电是不容易的。因此，传感器网络的生命周期依赖于传感器节点的电池能量。数据聚合技术通过减少在网络中的数据传输量大大提高了传感器节点的资源利用率。

（3）冗余　无限传感器网络的高度不可预知性和传感器节点的通信距离较短，因此需要一种高的节点冗余。传感器节点一般以高度的连通部署来解决传感器节点的故障问题。有了这种冗余，单个传感器节点的故障对于整个传感器网络能力的影响可以忽略。而且，因为单个传感器节点的数据可能产生误差，因此冗余信息用于支持服务质量和准确度。因此目标区域必须被传感器节点覆盖，这些传感器的感知范围有重叠。然后一个传感器节点的数据可以被感知相同事件的其他传感器节点的数据校正。然而，这样的信息冗余增加了从传感器节点到基站的传输数据量，从而大大减小了网络的生命周期。因此，高数据冗余必须由数据聚合技术来降低。

（4）敏感的安全性　许多无线传感器网络应用，例如监视、军事跟踪或生物医学，都是具有高敏感的安全性。因为资源的限制，不可能处理所有可能的安全问题，但无线传感器网络很容易受到节点捕获攻击。节点捕获攻击在传统的网络中不存在；因此传统网络的安全解决方案不能在无线传感器网络中使用。从数据聚合的观点来看，安全性是设计和发展高效的数据聚合协议的一大障碍。安全性协议要求传感器节点的数据在传输前被加密，而且更愿意数据被基站解密。然而，数据聚合协议更喜欢在每一个中间节点实现数据聚合。因此对于无限传感器网络安全性协议的设计必须考虑传感器节点的限制、节点捕获攻击和数据聚合。

（5）以数据为中心的处理　以数据为中心的处理是无线传感器网络的固有特性。应用程序对传感器节点的 ID 不感兴趣；因此在传感器网络中数据的命名方案通常是面向数据的。例如，一个环境检测系统通过查询获得温度的读数，查询是“收集矩形（x_1，y_1，x_2，y_2）区域范围内的温度读数”而不是“从一系列传感器 ID 为 x，y，

z的传感器收集温度读数。这样的以数据为中心的处理为数据聚合协议提供了良好的环境。无线传感器网络以数据为中心的处理特性支持数据聚合的过程。因此传输数据的传感器节点的ID号对于基站来说是不重要的，数据聚合协议能够很容易地结合和压缩采集到的数据。

（6）实时的限制　因为无线传感器节点处理现实世界中的数据，所以通信满足实时的约束是必要的。在边境监视系统中，例如在监测中的通信实验，在检测作用回路中的通信延迟会直接影响目标跟踪的质量。另一方面，因为大量的传感器节点、低成本的传感器硬件、恶劣的气候条件和不利的环境，传感器节点故障时常可见。同样，被密集部署的传感器节点共享的无线媒介受到阻塞和干扰。而且，高误码率、低带宽和非对称的信道使得通信难以预测。由于无线通信的性质和不可预测性，保证严格的实时约束是不可行的。数据聚合是影响实时约束的另一个因素，因为数据聚合增加了更多的传输时延。因此，数据聚合协议设计必须权衡能源利用率和实时约束。例如，在人的生命受到威胁的应用中，数据聚合可能不再使用，而是来满足应用的实时约束。

11.3　数据聚合

在典型的无线传感器网络中，大量的传感器节点从环境中采集应用查询所指定的信息，这些信息被传输到一个中心基站，它们在这里被应用程序处理、分析和使用。在这些资源受限的网络中，一般的处理方法是共同处理数据，这些数据是由传感器在向基站传输数据时产生的。这种网内的分布式处理数据通常被称为数据聚合，它包括结合属于同一现象或处理（也就是，计算平均值）的传感器数据。数据聚合的主要目标是通过减少传感器节点的资源消耗（例如电池能量和带宽）来延长网络的生命周期。在延长网络生命周期的同时，数据聚合协议可能会降低无线传感器网络中两个重要的服务质量指标：数据准确度和时延。因此，由于数据聚合协议的设计者必须权衡能源利用率、数据准确度和时延，所以设计一个高效的数据聚合协议是一项具有挑战性的任务。为了实现这个权衡，数据聚合技术与数据包如何通过网络路由紧密结合，因此，传感器网络的结构在不同的数据聚合协议中扮演着非常重要的角色。有一些协议允许路由和聚合的数据包同时存在。这些协议分为两类：基于树的数据聚合协议和基于分簇的数据聚合协议。以前有关数据聚合的研究关注于改善现有的路由算法来使数据聚合成为可能。结果，提出了许多基于最短路径树结构的数据聚合协议。为了减少由于基于树的数据聚合的时延，最近有关数据聚合的研究趋向于把传感器节点分组为簇，从而使数据的聚合在每一个组上进行以提高效率。除了这两类，还有基于多路径的数据聚合协议[4]。而且，还有一种基于树和多路径的混合方案。尽管，多路径和混合协议的效率不如基于树和分簇的数据聚合协议，但我们在这章的最后也会给出这些方案的简要概述。

11.3.1　基于树的数据聚合协议

最简单的实现分布式数据聚合的方法是在网络中确定一些数据聚合的节点，并保证传感器节点的数据路径包括这些数据聚合节点。这种基于树的数据聚合技术已经在参考文献［5－18］中被广泛研究。基于树的数据聚合协议的最主要问题是构造一棵高效的数据聚合树。如图 11-2 所示是一个基于树的数据聚合的例子。贪心增量树（Greedy Incremental Tree，GIT）是一个以数据为中心的路由协议，它允许基于定向扩展的数据聚合。在参考文献［7］中，GIT 与其他两种以数据为中心的路由协议比较，这两种协议分别被称为是最近来源的中心（Center at Nearest Source，CNS）和最短路径树（Shortest Path Tree，SPT）。仿真结果表明 GIT 就平均传输量而言是最好的。在参考文献［8］中提出了一个基于 SPT 的数据聚合协议，它可以促进父节点的节能意识。在这个协议当中，父节点的选择依赖于传感器节点到基站的距离和它们剩余的能量水平。也有一些数据聚合协议把信息理论作为路由标准。例如，在参考文献［9］中，提出了一个集中的方式，根据它们的联合熵来路由数据包。然而，这个协议对于无线传感器网络是不适用的，因为它依赖于每个传感器节点的消息熵和每个节点对的联合熵的全部知识。在本节的其余部分，我们将详细地展示在基于树的数据聚合中的一些重要工作。

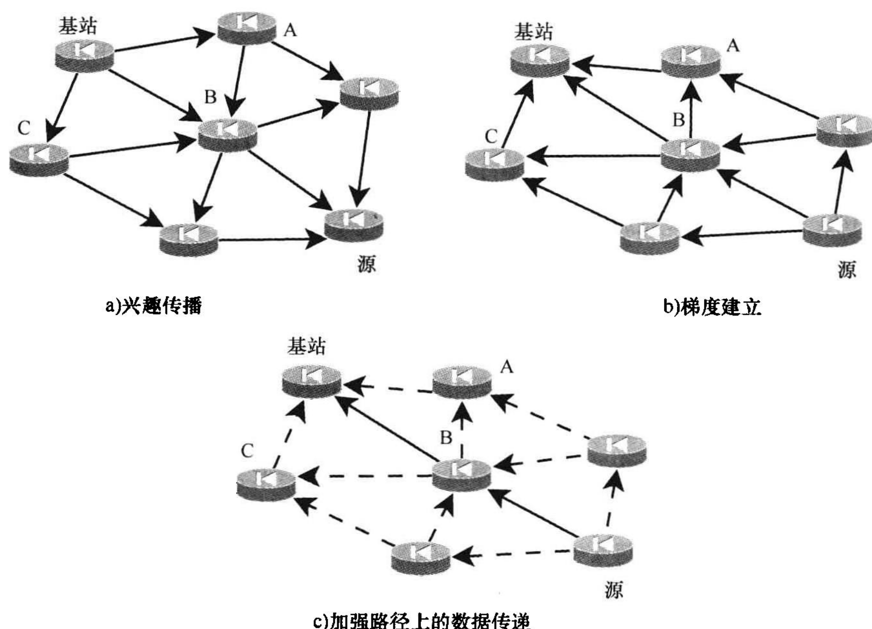


图 11-2　定向扩散的例子

在参考文献［10］中，Madden 等提出了一个以数据为中心的数据聚合框架，被

称为微小聚合（Tiny AGgregation，TAG）服务，它是基于SPT路由的。TAG是特意为监控应用程序设计的，它允许传感器节点调节睡眠时间表。为了实现这个，父节点让它们的子节点知道传输等待时间。同样，父节点缓存子节点的数据以防止数据丢失。TAG进行数据聚合分两阶段进行。第一阶段，称为分发阶段，面向基站的查询被分发到传感器节点，然后在第二阶段，称为收集阶段，聚合来自聚合路由的传感器读数。在分发阶段，基站广播一条消息，这一消息要求传感器节点组织一棵路由树，从而使基站能够发送它的查询。每条消息都有一个字段用于指定到（从）发送节点的根的层次或距离（根的层次为0)。当一个不属于任何层次的节点收到这一消息时，它将当前消息的层次加1，设置为自己的层次，并把发送者作为自己的父节点。这个过程将持续到在网络中的所有节点都加入到这棵树中，并有一个父节点为止。为了不断更新树结构，这条消息是周期性重复的。这条路由信息被sink节点定期地广播用于保持树结构的更新。一旦这棵树建好，基站通过这棵聚合树查询网络。当回复基站查询时，传感器节点会使用它们的父节点。TAG使用类SQL语言查询网络。每个查询需要指定查询的数量、聚合函数，需要进行数据收集的传感器节点。

定向扩散[11]是以数据为中心的协议，它发生在三个阶段：①信息传播；②梯度设置；③路径补充和转发。在第一阶段，基站传播一条兴趣消息，用于描述需要收集数据的类型和收集运作的模式。此消息一经接收，每个节点都向邻居节点广播。传感器节点准备好兴趣梯度，梯度基本上是向量，包括用于向基站回发查询结果的下一跳。对于每一种数据类型会设置不同的梯度。在设置梯度的最后，对于某些类型的数据，只有一个路径路由数据包到达sink节点（路径补充和转发)。图11-3给出了一个定向扩散协议的说明例子。数据聚合在数据转发阶段执行。基站定期刷新数据收集树，此树由补充路径组成。然而，这是一个昂贵的操作，如果这个网络有一个动态的拓扑结构，那么它可以克服由数据聚合产生的增益。在参考文献［12］中提出了一个定向扩散的修改版，称为增强定向扩散（Enhanced Directed Diffusion，EDD)，它集合具有基于分簇的结构的定向扩散，从而提高在梯度设置阶段本地交互的效率。在参考文献［13］中提出了另一个相似的协议。

在参考文献［14］中提出了传感器信息系统中的能量有效聚合（Power-Efficient GAthering in Sensor Information System，PEGASIS)，为了实现数据聚合它把传感器节点组织成一条链。每一条数据聚合链都有一个领导者，它用于向基站传输聚合数据。为了在网络中均匀的分配能量消耗，节点轮流做这条链的领导者。该链的实现采用在基站使用集中式方式或在每个节点使用贪心算法的方式。两种方法都需要网络的全部信息。该链的构造过程是从离基站最远的节点开始，并继续向基站推进。每次数据传输中节点在一条链中的位置移动一下，从而使错误最小化。在一条传感器节点链中，每个传感器节点收到来自邻居节点的数据，并和自己的数据聚合产生一个数据包，这个数据包和接收到的数据长度相同。这个过程在链中不断重复，这个链中的领导者将自己的数据添加到这个包中，并将它直接发送给基站。PEGASIS两个主要的缺点已被

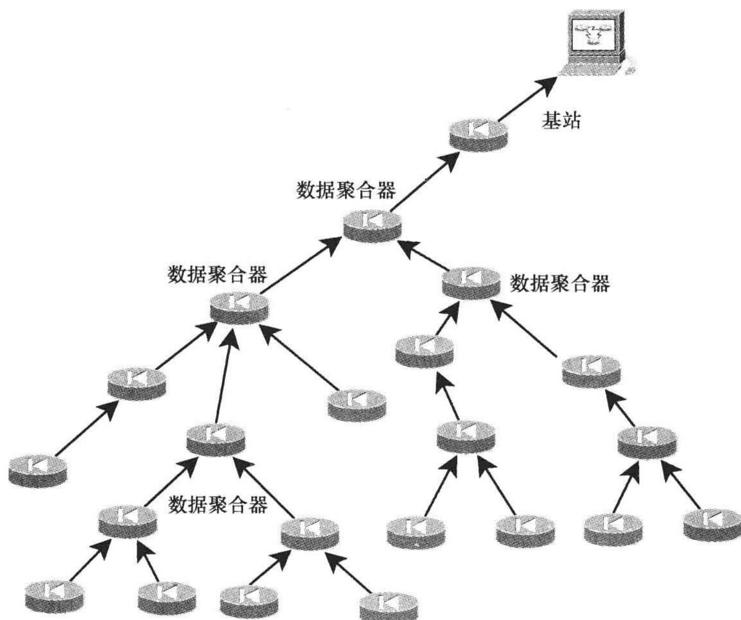


图 11-3 基于树的数据聚合

发现。第一，PEGASIS 要求每一个节点都有一个完整的网络拓扑视图，从而使链可以正确地形成。而且，所有的节点必须能够向基站直接发送数据。第二，如果在一条链中，传感器节点之间的距离过大，那么传感器节点的能量消耗将非常大。

在参考文献 [15] 中提出了一个数据聚合树的构造协议，它只依赖于网络拓扑的本地信息。该协议被称为 EADAT，它是基于能量感知的分布式的启发算法。基站就是聚合树的根节点，因此它通过广播一条控制消息来形成这棵树，这条消息包含以下五个部分：*ID* 号、父节点、能量、状态和跳数。这条消息在传感器节点间转发直到每个节点都广播了这条消息，其结果是形成一棵根节点为基站的聚合树。通过考虑传感器节点的能量水平，该算法中能量剩余较多的节点大多不会成为叶子节点，从而使数据转发任务由能量较高的节点来执行。仿真结果表明 EADAT 能够延长网络的生命周期，与不使用聚合的路由方法相比可以节省更多的能量。还发现与不使用数据聚合的情况相比，传感器节点的平均能耗水平下降得非常缓慢。因此，当最大化网络的生命周期非常必要时，EADAT 可以用于构造节能高效的数据聚合树。

在无线传感器网络中还有很多其他的方法用于解决构造高效数据聚合树的问题。在参考文献 [16] 中提出了一种不同的方法，称为时延限制介质访问控制（Delay Bounded Medium Access Control，DBMAC），它集成了路由和 MAC 协议进行数据聚合。提出 DBMAC 解决方案的主要目标是最小化时延和通过利用数据聚合的机制提高能源利用率。DBMAC 采用载波监听多址接入/冲突避免（CSMA/CA）媒体访问模式基于请求发送/清除发送/数据/确认（RTS/CTS/DATA/ACK）握手。通过利用其他节点的

CTS 消息，传感器节点可以在队列中已经有一些数据包要传输的节点中选择，作为中继节点。这个过程提高了网络中数据聚合的效率，因为所有沿路径存储的信息被聚集到一棵树。通过显示从跨层设计上获得能量有效的数据聚合策略，DBMAC 被认为是路由和数据聚合如何相互影响的很好的例子。在另一个参考文献［17］中，提出了基于树的动态协作（Dynamic Convoy Tree-Based Collaboration，DCTC）。DCTC 的目标在于通过平衡目标区域的聚合树来减少能量消耗。然而，DCTC 带来了沉重的消息交换，并假设传感器节点具有到达中心事件的距离信息，这些距离信息在所有的应用程序中是不能通过使用检测到的信息计算得到的。在参考文献［18］中，提出了一个高效的数据收集和聚合的协议（Power-Efficient Data Gathering and Aggregation Protocol，PEDAP），就数据传输的数量而言，它用于最大化网络的生命周期。在该协议中，每一轮对应数据的聚合，这些数据从不同传感器节点传输到基站。PEDAP 是以最小生成树为基础的协议，当基站位于目标区域内时，比 LEACH 和 PEGASIS 等协议具有更好的表现。

11.3.2 基于分簇的数据聚合协议

在基于分簇的数据聚合协议中，传感器节点被分为簇。在每一个簇中，选择簇头来聚合局部的数据和传输聚合结果给基站。簇头可以通过远距离无线传输和 sink 节点直接通信；但是这对于能量受限的传感器节点来说是非常低效的。因此，簇头通常形成一个树结构通过经过其他簇头的多跳来传输聚合的数据，这样有助于节约能源。图 11-4 给出了一个基于分簇的数据聚合的例子。最近，在参考文献［19-25］中提出了一些基于分簇的数据聚合协议。

在参考文献［19］中，提出了一个自组织和自适应的分簇协议，称为低能量自适应的分簇等级协议（Low-Energy Adaptive Clustering Hierarchy，LEACH）。LEACH 利用随机性来均匀分配节点之间的能量消耗。LEACH 是一种分簇的方法，在这方法中簇头作为数据汇聚点。这个协议包括两个阶段。在第一个阶段，簇结构形成，在第二个阶段，簇头聚集并传输数据到基站。LEACH 的簇头的选择过程依赖于概率分布的方法。在每个数据聚合选择阶段，传感器节点计算以下门限。

$$T(n)=\begin{cases}\dfrac{P}{1-P(R_{\text{mod}}(1-P))} & n\in G\\ 0 & \text{其他}\end{cases}$$

式中 P——要求的簇头所占的百分比；

R——轮数；

G——在最后 $1/P$ 轮期间还未成为簇头的节点集。

为了成为簇头，一个传感器节点在［0，1］之间随机选取一个随机数，如果选择的数小于 $T(n)$，那么这个节点成为簇头。簇头公告被广播到传感器节点，传感器

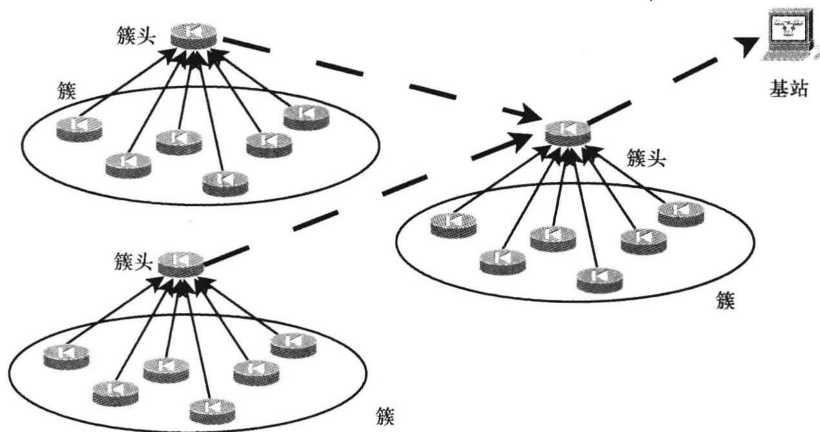


图 11-4　基于分簇的数据聚合

节点根据公告消息的信号强度来加入簇。根据簇成员的数量，每个簇头根据 TDMA 来调度它的簇成员，从而达到局部传输的最优化管理。在第二阶段，传感器节点根据已经建立好的时间表发送数据到簇头。通常传感器节点会随意地关闭无线电，直到它们预定的 TDMA 传输时间段到达。LEACH 要求簇头通过单一链接向基站发送它们的数据。然而，这正是 LEACH 协议的缺点，因为基站离簇头很远时单一链接传输是非常昂贵的。因为 LEACH 完全不需要有关网络结构的信息，所以它是完全分布式的。就簇头的选择而言，它也是自适应的。另一方面，如果由于移动的节点网络拓扑是动态的，那么可能会有很高的控制消息开销。在参考文献［20］中，提出了另一个基于分簇的数据聚合协议，称为 HEED。对于簇头的选择，HEED 得益于传感器节点多种能量等级的可用性。事实上，一种联合的公制是由节点的剩余能量和节点到邻居节点的距离组成。HEED 定义最小能量等级的平均值是所有在簇内部的传感器到达簇头所需的能量。这就是所谓的平均最低可达能量（Average Minimum Reachability Power, AMRP）。AMRP 用于计算每个簇的通信开销。为了选择簇头，每个传感器节点计算它选为簇头的概率，如下所示。

$$P_{(\mathrm{CH})} = C \times \frac{E_{\mathrm{residual}}}{E_{\max}}$$

这里初始簇头的百分比记为 C，E_{residual} 和 $E_{\max}$ 分别代表传感器节点当前的剩余能量和初始的能量。每一个节点广播簇头消息；当在收到群头消息的节点中，其中一个节点的 AMRP 最小，那么传感器节点选择它作为它们的群头。这个过程重复地进行，直到所有的节点都被分配到一个群头。与 LEACH 相似，在 HEED 协议中，群头直接和基站通信。仿真结果显示 HEED 延长了网络的生命周期，并使群头集在地域上均衡。

在参考文献［21］中，提出了一个每一跳定期进行数据聚合的分簇方案，该方案称为 Cougar，它对于传感器节点持续生成相关数据的应用是合适的。一旦簇头聚合它们

的簇内数据，它们就发送本地的聚合数据到网关节点。与LEACH相似，对于动态的网络拓扑，Cougar存在负面的影响，然而Cougar拥有一个独特的簇头选择过程。Cougar根据多于一个标准选择簇头，并允许传感器节点能够多跳远离它们的簇头。这就要求路由算法交换群中的数据包。Cougar使用按需距离向量自组织（Ad Hoc On Demand Distance Vector，AODV）路由协议选择簇内中继。在Cougar协议中，采用同步机制来正确地聚合数据。在簇中，簇头和所有的传感器节点都是同步的，直到所有节点都传输了它们的数据，它才将聚合数据报告给网关节点。

分簇扩散和动态数据聚合（Clustered Diffusion with Dynamic Data Aggregation，CLUDDA）[22]是一种混合的方法，它使用扩散机制结合簇。CLUDDA包括查询定义，由基站发起感兴趣的消息。每一个感兴趣的消息包括查询的定义，定义描述了需要在数据组件执行来产生适当回应的操作。感兴趣的消息的传输通过利用现有的查询知识来减少处理开销。CLUDDA在兴趣消息广播的发起阶段将定向扩散[11]和分簇结合起来。通过使用分簇机制，它保证了只有进行簇间通信的簇头参与感兴趣消息的传输。因为一般的传感器节点不需要传输任何数据，除非它们有能力满足服务要求，所以CLUDDA能够节约能源。在CLUDDA协议中，任何具有查询定义的消息的簇头能够进行数据聚合，因此聚合点是动态的。同样，每一个簇头维护一个查询缓存来显示不同的数据组件，这些数据组件被聚合来获得最终的数据。簇头还保存了邻居节点的地址的列表，数据信息来自邻居节点。这些地址用于直接向指定的节点传送感兴趣的消息，而不是采用广播。

在参考文献中还提出了一些其他的基于分簇的数据聚合协议。它们中的一些是对现有协议的改善。在参考文献［24］中，采用一种跨层的方法将MAC设计集成到数据聚合的概念中。在参考文献［25］中提出了一个基于位置的分簇方案，在该方案中传感器节点自组织形成静态簇。传感器节点沿最短路径向簇头发送它们的数据，网络内部的聚合在节点之间执行。簇头通过多路径执行数据聚合并向基站发送聚合数据。然而，在簇头向基站传送聚合数据期间没有进一步地执行数据聚合。

11.3.3 基于多路径的数据聚合协议

除了基于树和分簇的数据聚合协议，同样还有基于多路径的数据聚合协议[26]，在这种协议中，传感器节点把它们聚合的数据分成几部分，并通过多路径向父节点发送这些数据分片。这种方案的主要思想是通过多路径向sink节点发送部分小数据块的复本，从而提高网络的稳定性。然而，因为发送复本数据，与基于树和分簇的协议相比，它们增加了通信开销。基于多路径的数据聚合协议通常使用环形拓扑，在环形拓扑中，传感器节点就跳数而言根据到基站的距离被分成几个等级。例如在概要扩散[26]中，当数据包一层一层向基站移动时，数据聚合通过多路径执行。基于树和多路径的方法都有它们自己的缺点，例如通信开销或链路故障。因此，在参考文献［27］中提出的方案目的在于通过结合这两个方案的最佳功能来解决这

两个方案的问题。该方案是一个混合的协议，称为 Tributaries and Deltas。在 Tributaries and Deltas 协议的数据聚合结构可能在网络的不同部分运行。Tributaries and Deltas 背后的动机是当丢包率低时采用数据聚合树，在丢包率高时利用多路径方案。为了实现这个目标，Tributaries and Deltas 把传感器分为两类：使用基于树的方法传输数据的节点和使用多路径的节点。不同区域的数据使用一些修正规则结合。为了总结本节，在图 11-5 中我们给出了数据聚合协议的比较。

	TAG[10]	D.Diffusion[11]	PEGASIS[14]	DBMAC[16]	LEACH[19]	COUGAR[21]	S.Diffusion[26]	T.and Deltas[27]
聚合方式	树	树	链状	分布式的	分簇的	分簇的	多路径	多路径/树
维护代价	高	高	高	低	中等的	中等的	中等的	中等的
移动支持	低	中等的	低	高	低	低	高	中等的
链路失败支持	中等的	中等的	低	中等的	低	中等的	高	高

图 11-5　数据聚合协议比较

11.4　安全的数据聚合

传感器网络资源受限的性质对于网络的安全性造成了很大的挑战。不仅是军事用途，在房屋和地基检测，建筑物检测、报警器和关键系统，例如机场和医院的安全性也是严格的。然而，在安全性和数据聚合协议之间存在强大的冲突。安全机制要求传感器节点在传输前对所有感知数据进行加密。而且，为了尽可能地避免安全问题，要求在基站采用对数据解密的端到端的安全机制。然而，数据聚合协议要求中间节点处理数据包，找出多余的需要进行加密的数据包。这两个冲突的目标要求数据聚合算法的设计考虑安全的通信算法。

无线传感器网络的安全要求可以通过对称密钥加密或非对称密钥加密得到满足。由于传感器节点的资源限制，在无线传感器网络中对称的密钥加密比非对称的密钥加密更合适。通过使用对称的密钥加密算法，保密和数据聚合可以在一种逐跳方式下一起实现。然而，在这种情况下，数据聚合必须对每一条收到的消息加密，并根据相应的聚合函数聚合消息，在传输数据前加密聚合的结果。因此，使用 hop-by-hop 加密算法，要实现终端到终端的保密是不可能的。而且，这个方案要求聚合和转发节点建立与它们直接相邻的节点的密钥。通过使用对称密钥加密算法实现数据聚合和安全的必要性使得许多研究者关注安全数据聚合问题[28-36]。除了参考文献［30］和［40］的所有方案中提到，数据聚合节点为数据聚合解密感知数据。最近，一系列数据聚合协议被提出，用于实现不需要传感器数据聚合时解密的数据聚合[37,40,43,44]。这些协议中的一些使用非对称的密钥加密算法，它们对于资源受限的传感器节点是适用的。不需要数据加密的数据聚合协议的缺点是它们仅适用于一部分聚合函数。接下来，我们在两小节展示安全数据聚合协议：执行普通数据上数

据聚合的安全数据聚合协议，和使用加密数据聚合的安全数据聚合协议。

11.4.1 在普通的数据上的安全数据聚合

早期的安全数据聚合集中在普通数据的聚合。在参考文献［28］中，这个安全机制检查节点的不良行为，例如丢弃或伪造信息和传输错误的数据。在参考文献［29］中，随机抽样机制和交互证明用于在基站检查聚合数据的准确性。在参考文献［31］中，聚合数据的执行节点同样聚合数据，并计算MAC帮助证明基站的聚合数据的准确性。因为数据验证在基站执行，所以错误数据的传输和MAC注册到基站严重地影响传感器网络资源的利用率。在参考文献［32］中，只有到一个错误的行为被检测到，传感器节点才使用加密算法。拓扑的限制引入了一棵促进数据聚合监视的安全聚合树。在SAT中，任何子节点都能监听到父节点的输入数据。当一个数据聚合的聚合数据可疑时，使用一个加权表决方案来决定数据聚合是正确的行为还是欺骗的行为。在参考文献［35］中，提出了一个安全的逐跳的数据聚合协议（SDAP）。SDAP的创始人被这样的事实所激励，在正常的逐跳的树拓扑聚合过程中，与低等级的传感器节点相比，高等级的节点（也就是更靠近根节点的数据）更可靠。这是因为聚合数据被一个高层次的节点计算，被高层次节点计算的聚合数据代表了大量低层次的传感器节点的数据。如果一个被攻占的节点离基站更近，那么这个节点会产生错误的结果，这对基站的最后聚合结果会有很大的影响。因为所有的传感器节点有容易被攻占的简单硬件，其他节点比这些成本低的传感器节点更值得信任。因此，SDAP的目标在于通过分而治之的原则增加该方法中对高层次节点的信任。SDAP利用概率的方法把拓扑树动态的分成多个大小相同的逻辑组。用这种方法，在一个逻辑组中更少的节点被分布在高层次的传感器节点，从而减少由一个被攻破的高层次节点导致的潜在的安全问题。图11-6展示了一个分组树的例子。

在参考文献［36］中，作者认为单独的加密基元不能够提供足以解决安全数据聚合问题的方案，因为被攻破的节点可以访问用来保护聚合过程的密钥。基于这个观察，作者提出了一个安全可靠的数据聚合协议，称为SELDA，它是基于传感器节点的可靠性和数据聚合的，利用网络信任来克服基于加密的安全数据聚合方法的缺点。作为测试分布函数族[41]能够用来预测传感器节点动作的后验概率[42]，在SELDA协议中通过传感器节点不良行为的测试分布函数来估计传感器节点的可靠性。SELDA协议的基本思想是传感器节点观察它们邻近节点的行为来建立对环境和邻近节点的信任级别。传感器节点使用检测的机制来检测节点的可靠性，感知和路由以及邻居节点的不良行为。这些不良行为将测试分布函数作为量化信任级别准。图11-7说明了在SELDA协议中不良行为的检测。传感器节点通过邻居节点交换它们的信任级别来形成信任网络，确定安全可靠的数据聚合路径。基于这些信任级别，传感器节点通过一个或多个安全路径传输它们的数据。在数据聚合期间，数

据聚合根据发送数据的信任级别来加权数据。仿真结果表明 SELDA 协议以一个可以容忍的通信开销提高了数据聚合的可靠性。

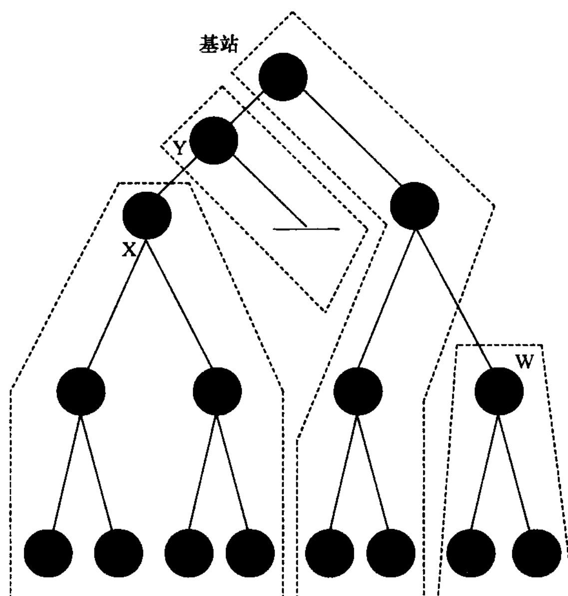


图 11-6　在 SDAP 中一个聚合树的例子。颜色为黑的节点 X、Y、W 是领导节点，基站作为根是默认的领导。

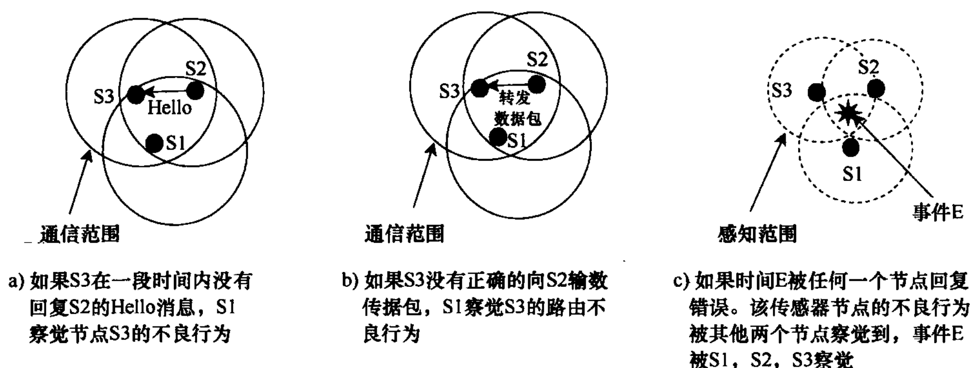


图 11-7　在 SELDA 协议中不良行为的检测

上面的所有协议都使用真实的感知数据来进行数据聚合，因此聚合时需要对感知数据进行解密。在本节的其余部分，我们会呈现两个另外的数据聚合协议，它们不需要使用真实的数据，因此安全和数据聚合可以被一起实现。

在参考文献［30］中，作者提出了一种能量高效的和基于安全模式的数据聚合协议（Energy-efficient and Secure Pattern-based Data Aggregation，ESPDA），在基于分簇的无线传感器网络中，它同时考虑了数据聚合和安全这两个概念。ESPDA

是首次研究考虑数据聚合技术而不危及安全的，它使用模式码来执行数据聚合。该模式码基本上是具有代表性的数据项目，根据每一个模式码都有相应的真实数据的某些特征，从真实数据中提取，采用这种方式每个模式码有某些特征是与真实数据相对应的。根据真实数据的具体类型，提取过程可能是多样的。例如，当真实数据是由监视传感器感应的人类的图像时，根据应用程序的要求脸部和身体的关键参数被认为是代表性数据。当一个传感器包括多个感知单元，通过结合单个感知单元的模式码来获得传感器节点的模式码。代替了传输整个感知数据，传感器节点生成并传输模式码给簇头。簇头决定了不同的模式码，然后只要求一个传感器节点为每个不同的模式码传输真实的数据。这个方法使 ESPDA 能源和带宽都是高效的。因为簇头无须为数据聚合的数据加密，没有加密密钥或解密密钥被广播，因此 ESPDA 协议同样是安全的。此外，非阻塞 OVSF 块跳频技术通过随机改变数据块映射到 NOVSF 时间段进一步提高 ESPDA 的安全性。ESPDA 假设传感器节点是以高密度分布的，导致节点的感知区域高度部分重叠，而引起的传感器节点失效问题。网络的随机部署使得许多区域被多个传感器节点覆盖。因此，强烈要求它能够保证任何时刻一个区域只被一个传感器覆盖，因此只有一个传感器会感知相同的数据。因为冗余的数据无法被感知，因此提高了数据聚合效率。

在 ESPDA 协议中，传感器节点被随机分布在目标区域进行监控，在初始化部署之后，把它们组织成簇。从每一簇中选择一个簇头来处理群节点和基站之间的通信。为了使所有节点间具有统一的电源消耗，根据剩余能量来动态改变簇头。因为数据传输是能量消耗的重要原因，ESPDA 在睡眠-工作模式协调协议的帮助之下，首先减少传感器节点到簇头冗余数据的传输。然后，为节约能量最小化传输数据量，为了节约能量，数据聚合用来消除冗余和最小化传输数据量。在传统的数据聚合方法中，簇头节点接收来自传感器节点的所有数据，然后通过检查传感器数据的类型来消除冗余。ESPDA 使用模式码来代替感知数据执行数据聚合。因此，在簇头（节点）处不需要知道传输数据的类型（内容）。这使得 ESPDA 能够和安全协议协调工作。在安全协议中，传感器数据被簇头节点定义为无冗余的，它以加密的形式被传输到基站。模式码通过使用簇头节点周期性广播的秘密模式种子来产生。模式种子始终都不允许有相同的模式码产生，其是用于提高模式码保密性的随机数。当模式种子改变，模式产生算法为相同的传感器数据生成一个不同的模式码。因此，在感知数据被传感器节点传输前要消除冗余。

在基于分簇的无线传感器网络中，根据结合数据聚合和安全概念的研究，提出了安全的基于参考的数据聚合协议（Secure Reference-Based Data Aggreation，SRDA）。与 ESPDA 相似，SRDA 同样意识到数据聚合协议要在数据通信安全协议下协调工作，因为它们之间的任何冲突都会对网络的安全产生负面的影响。在 SRDA 协议中，将传感器节点感知的原始数据和参考数据值进行了比较，然后只传输差值数据。参考数据被认为是大量以前的传感器读数的平均值。SRDA 背后的动机是由于

无线电通信是传感器节点能量消耗的主要原因，减少传输的比特数可以节省大量能量。尽管数据聚合减少了数据包的数量，但增大了传输数据包的大小，这将会减少能源消耗。在传统的数据聚合算法中，传感器节点将原始的数据传输给它们的簇头。因为在每个数据包中一定范围的数据可能仍是相同的，这将引起能量和带宽的浪费。然而，SRDA 传输不同的数据而不是原始的感知数据。也就是说，将传感器节点感知的原始数据和参考数据相比，然后只传输差值数据。例如，用 102℉表示一个传感器节点的温度测量值。如果 100℉被簇头认为是参考温度，那么传感器节点只传输在传输过程中参考值和当前测量值的差异（也就是 2℉）。结果，差异的聚合有减少传感器节点到簇头节点传输数据量的巨大潜力。差异数据聚合背后的基本动机是只有当在环境中发生重要的事件（例如对于一个温度检测网络来说发生火灾事件），传感器测量才发生巨大的变化。通常在传感器网络中，这些所谓的重要事件比普通事件的发生率低。

SRDA 在数据传输会话层实现，会话是指从目前建立传感器节点和簇头之间的通信到通信终止这段时间间隔。每一个会话预计将有大量的数据包。SRDA 是独立的分簇方案，它适用于任何分簇算法。因为与无线传感器节点范围较低的传输相比，接收同样消耗大量能量，所以接收和发送节点都从这种技术收益。基于参考的数据聚合协议适用于分簇层次的各级。由于原始数据不通过包传输，因此当参考值大于差异值时，这种技术的效率更高。这种技术执行的另一个重要因素是对连续数据包内容值的方差，由于较小的数位就能够代表不同的数据值，因而方差越小通过差异数据聚合实现的获利将增加。

11.4.2　对加密数据的安全数据聚合

使用传统的对称数据加密算法，端到端的保密性和数据聚合不能一起实现。如果对称密钥的加密算法的应用与高效的数据聚合相结合，然后消息必须逐跳加密。但是，这意味着，为了执行数据聚合，中间节点必须对每一条收到的消息解密，然后根据相应的聚合函数来聚合消息，最后在传输之前对聚合结果加密。而且，这个过程要求邻近的数据聚合共享加密和解密的密钥。为了不需要在数据聚合间共享密钥而一起实现终端到终端的数据保密和数据聚合，隐私同态加密已在参考文献[37，40，43，44] 中使用。

隐私同态是一种加密的转变，它允许对加密的数据直接进行计算。设 E 为加密，D 为解密。同样，在 Q 数据集合上，设 + 表示加法，× 表示乘法。假设 K_{pr} 和 K_{pu} 分别是基站的私钥和公钥。加密转换被认为是加法的同态，如果

$$a+b=D_{K_{pr}}(E_{K_{pu}}(a)+E_{K_{pu}}(b)) \qquad a,b\in Q$$

它接受的是乘法同态，如果

$$a+b=D_{K_{pr}}(E_{K_{pu}}(a)\times E_{K_{pu}}(b)) \qquad a,b\in Q$$

因为加法和乘法同态加密函数分别支持对加密数据进行加法和乘法操作，数据聚合可以通过对加密的数据执行基于加法和乘法的数据聚合。

在CDA[37]中，传感器节点和基站共享一个公共的密钥，对来自中间路径的数据聚合保持隐藏。这项工作的重要贡献是为传感器节点和基站之间的反向多重广播提供端到端的加密。在该方法中，数据聚合执行适用于密文（加密数据）的聚合函数。这提供了优势，使中间聚合不需要执行高消耗的加密和解密操作。因此，数据聚合不需要存储敏感的密钥，它保证了在无线传感器网络的生命周期中一个无限制的数据聚合节点的选择过程。因为只有那些存储密钥的节点才能作为数据聚合节点，因此对于逐跳的加密无限制的数据聚合选择是不可能的。作为隐私同态加密函数，该协议利用Domingo-Ferrer提出的函数[38]。Domingo-Ferrer的加密函数是概率的，因为加密变换在从一系列可能的密文中选择相应的给定的密文具有随机性。

Domingo-Ferrer的加密函数的公共参数是$d \geqslant 2$的正整数和一个必须要有很多小除数的超大整数g。而且，应该要有许多小于g的整数，它们可以被g倒模。密钥通过$k=(r, g')$计算。$r \in Z_g$，由$r^{-1} \bmod g$选择，这里$\lg_{g'} g$表示由该函数提供的安全级别。明文集是$Z_{g'}$，密文集是$(Z_g)^d$。加密和解密的过程定义如下。

加密：把$a \in Z_{g'}$随机分裂成$a_1 \cdots a_d$，使得$\sum_{j=1}^{d}(a_j \bmod g')$，并且$a \in Z_{g'}$计算

$$E_k(a) = (a_1 r^1 \bmod g, a_2 r^2 \bmod g, \cdots, a_d r^d \bmod g)$$

解密：通过$r^{-j} \operatorname{mog} g$计算第$j$个坐标重新获得$a_j \bmod g$。为了获得$a$，计算如下。

$$D_k(E_k(a)) = \sum_{j=1}^{d}(a_j \bmod g')$$

密文×操作由Z_g中的所有交叉相乘执行。然后把具有相同程度的条款加起来。密文+操作比×操作相对简单，且是分量执行。

从上面的定义来看，Domingo-Ferrer的非对称的基于密钥的隐私同态对于资源受限的传感器节点来说计算代价是昂贵的。参考文献［37］的作者比较非对称的基于密钥的隐私同态和对称的基于密钥的隐私同态所需的时钟周期。结果显示加密、解密和加法对于执行Domingo-Ferrer的函数是必要的，与那些执行所必需的对称的基于密钥的RC5相比更加昂贵。然而，作者认为因为CDA可以有效地平衡能量消耗，因此它的这个缺点是可接受的。使用对称的基于密钥的加密方法执行一跳一跳的数据聚合导致数据聚合节点的生命周期变短。因此，当维护一个互连的无线传感器网络的骨干时，数据聚合称为执行的瓶颈，这时使用CDA的非对称的基于密钥的隐私同态来平衡数据聚合的能量消耗是较合适的。

在参考文献［43］中，一个安全数据聚合协议，称为CDAP，它利用非对称的基于密钥的隐私同态的加密方法来实现端到端的数据保密和数据聚合。作者指出非对称的基于密钥的隐私同态带来更高的计算消耗；那些资源受限的节点无法承担起网络的加密和聚合开销。因此，对于隐私同态加密和加密数据的聚合，CDAP使用

一系列资源丰富的传感器节点，被称为聚合节点（Aggregator Node，AGGNODE），在 CDAP 协议中，在网络部署之后每一个 AGGNODE 建立和它的邻近节点的成对密钥，从而使邻居节点可以安全地发送它们的读数到 AGGNODE。在 CDAP 协议的数据收集阶段，每个 AGGNODE 查询其邻居节点的传感器读数，例如温度或湿度。每一个邻居节点通过使用对称密钥加密算法（RC5）加密它的数据，并把它发送到 AGGNODE。AGGNODE 解密从邻居节点收到的所有数据，并聚合它们，然后通过使用隐私同态加密算法对聚合的数据进行加密。一旦数据被隐私同态加密算法加密，只有基站可以通过使用它的私有密钥进行解密。然而，因为同态的性质，尽管中间的 AGGNODE 没有基站的私有的密钥，也可以聚合这节加密的数据。因此，由传感器节点收集的数据在传输到基站的过程中是被 AGGNODE 聚合的。基站通过使用它私有的密钥解密最终的聚合数据。在图 11-8 给出了一个说明 CDAP 的例子。因为隐私同态加密算法的计算开销，在 CDAP 中，只有 AGGNODE 可以使用隐私同态加密算法来加密和聚合收集到的数据。

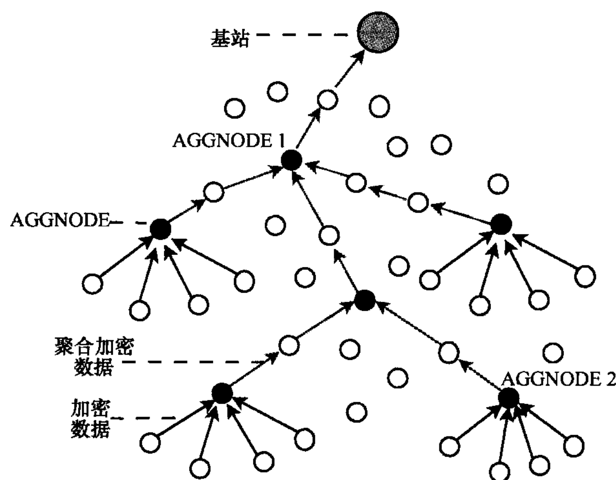


图 11-8　显示了 CDAP 协议的聚合情况。当数据传输到基站时，AGGNODE 收集来自它们的邻居节点的数据，并加密被 AGGNODE 聚合的数据

因此，在 CDAP 协议的数据收集第一阶段，传感器节点使用对称的密钥算法进行加密。因为对称的加密，一个被攻占的 AGGNODE 会泄露它邻居节点数据的密钥或往数据中注入错误的数据。然而，作者认为这种攻击的影响是局部的，是可以接受的。

参考文献［44］提出了一个简单的和可证明是安全的加法同态的流密码，它允许加密数据的高效聚合。该技术基于使用对 n 取模的加法操作的一次性加密技术的延伸。这个方案的主要思想是用模加操作（+）代替流密码的独占异或操作 *OR*。加密和解密操作可总结如下。把消息 m 表示为整数 $m \in [0, M-1]$，这里 M 是一个超大整数。同样，假设 k 是由密钥流产生的一个随机数，$k \in [0, M-1]$。然

后，密文 c 计算如下：$c = enc(m,k,M) = m + k(\mathrm{mod}M)$。为解密密文 c，执行 $\mathrm{Dec}(c,k,M) = c - k(\mathrm{mod}M)$。根据这些函数，密文被添加如下：使 $c_1 = \mathrm{Enc}(m_1, k_1 M)$，$c_2 = \mathrm{Enc}(m_2, k_2, M)$，然后，对于 $k = k_1 + k_2$，Dec（$c_1 + c_2$，k，M）$= m_1 + m_2$。假设消息 m，$0 \leqslant m < M$，因为除了拥有交换的性质，该方法是加法同态。由于这样的加密过程，该方案大大减少了传感器节点能量的消耗。然而，在这个方法中，由于节点或通信故障，每一个聚合消息都与对聚合没有用处的节点列表联系连接在一起。在参考文献［40］中，通过采用一个分层数据聚合模型，这个问题可以解决。为了总结这一节，在图11-9，我们给出了安全数据聚合协议的比较。

	数据保密	数据完整	源认证	可用性
Hu et al.[28]		●	●	
SIA.[29]	●	●	●	
ESPDA[30]	●	●	●	
Du et al.[31]		●	●	
Wu et al.[32]		●	●	
SRDA[34]	●	●	●	
SDAP[35]	●	●	●	
SELDA[36]		●	●	●
CDA[37]	●			
Ozdemir[40]	●			
CDAP[43]	●			
Castellucia et al.[44]	●			

图11-9　安全数据聚合协议的比较

11.5　开发性的研究问题和未来研究方向

这一章给出了在一个无线传感器网络中数据聚合的全面概述。介绍了目前最先进的数据聚合协议，根据网络拓扑和安全对它们进行分类。尽管提出的研究解决了数据聚合的许多问题，但还有很多研究领域需要和数据聚合过程相联系。

因为路由机制和数据聚合协议是相关的话题，所以它们之间的关系已经得到很好的研究。除了基于分发和树的数据聚合协议，提出了许多基于分簇的聚合协议，它们通过簇头来路由聚合数据。尽管，这些协议在簇结构长时间不变的静态网络中是非常高效的，但是在动态网络中它们的效率相当差。因此，在动态环境中的数据聚合将成为未来的一个研究方向。而且，到目前为止，数据聚合的信源编码理论的应用程序已经得到了不少关注。考虑到传感器节点是高度相关的，数据聚合可以通过使用信源编码技术来实现。目前在这方面的研究只注重理论成果，对于无线传感器网络还没有实际可用的算法。因此，对于今后在基于信源编码的数据聚合的研究

有很大的余地。在数据聚合协议中安全是另一个重要的问题。尽管在这方面许多协议已经被提出，但还有许多没有解决的问题，例如当在数据聚合时注入错误数据的被攻占的数据聚合，因为数据聚合经常导致数据中的改变，被攻占的数据聚合注入的错误数据很难被察觉到。检测被攻占数据聚合注入的错误数据将成为未来一个令人感兴趣的研究方向。数据聚合协议的传感器异构性的影响将是另一个未开发的研究领域。使用功能强大的传感器作为数据聚合器的协议展现出了可喜的成果。然而，为产生最好的数据聚合结果，这些强大的传感器位置的确定仍需进一步的研究。

11.6 总结

这一章给出了在无线传感器网络中关于数据聚合概念的详细介绍。为了给出数据聚合背后的动机，首先解释了无线传感器网络的独特性质，例如资源受限和以数据为中心的处理过程，这些都与数据聚合概念相关。然后，给出了目前最先进的数据聚合协议。安全和数据聚合之间的权衡被详细地研究，对安全数据聚合协议解释得也很详细。最后，我们给出了在数据聚合概念下开放性的研究问题和未来的研究方向作为本章的总结。

参考文献

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, A survey on sensor networks, *IEEE Communications Magazine*, 40(8), 102–114, Aug. 2002.
2. K. Akkaya, M. Demirbas, and R. S. Aygun, The impact of data aggregation on the performance of wireless sensor networks, *Wiley Wireless Communications and Mobile Computing (WCMC) Journal*, 8, 171–193, 2008.
3. Crossbow Technologies, Inc., http://www.xbow.com
4. E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, *IEEE Wireless Communications*, 14(2), 70–87 Apr. 2007.
5. C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, Impact of network density on data aggregation in wireless sensor networks, in *Proc. of 22nd International Conference on Distributed Computing Systems*, pp. 457–458, Vienna, Austria, Jul. 2002.
6. C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, Directed diffusion for wireless sensor networking, *IEEE/ACM Transactions on Networking*, 11(1), 2–16, Feb. 2003.
7. B. Krishnamachari, D. Estrin, and S. Wicker, The impact of data aggregation in wireless sensor networks, in *Proc. of 22nd International Conference on Distributed Computing Systems Workshops*, pp. 575–578, Vienna, Austria, Jul. 2002.
8. M. Ding, X. Cheng, and G. Xue, Aggregation tree construction in sensor networks, in *Proc. of the 58th IEEE Vehicular Technology Conference*, vol. 4, pp. 2168–2172, Oct. 2003.
9. R. Cristescu, B. Beferull-Lozano, and M. Vetterli, On network correlated data

gathering, in *Proc. of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2571–2582, Mar. 2004.
10. S. Madden et al., TAG: A tiny aggregation service for ad hoc sensor networks, in *OSDI 2002*, Boston, MA, Dec. 2002.
11. C. Intanagonwiwat et al., Directed diffusion for wireless sensor networking, *IEEE/ACM Trans. Netw.*, 11(1), 2–16, Feb. 2002.
12. B. Zhou et al., A hierarchical scheme for data aggregation in sensor network, in *IEEE ICON 04*, Singapore, Nov. 2004.
13. M. Lee and V. W. S. Wong, An energy-aware spanning tree algorithm for data aggregation in wireless sensor networks, in *IEEE PacRrim 2005*, Victoria, British Columbia, Canada, Aug. 2005.
14. S. Lindsey, C. Raghavendra, and K. M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, *IEEE Trans. Parallel Distrib. Sys.*, 13(9), 924–935, Sept. 2002.
15. M. Ding, X. Cheng, and G. Xue, Aggregation tree construction in sensor networks, in *IEEE VTC 03*, Orlando, FL, vol. 4, pp. 2168–2172, Oct. 2003.
16. G. Di Bacco, T. Melodia, and F. Cuomo, A MAC protocol for delay-bounded applications in wireless sensor networks, in *Med-Hoc-Net 2004*, Bodrum, Turkey, June 2004.
17. W. Zhang and G. Cao, DCTC: Dynamic convoy tree-based collaboration for target tracking in sensor networks, *IEEE Transactions on Wireless Communications*, 3(5), 1689–1701, Sep. 2004.
18. H. O. Tan and I. Korpeoglu, Power efficient data gathering and aggregation in wireless sensor networks, *SIGMOD Record*, 32(4), 66–71, Dec. 2003.
19. W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, *IEEE Trans. Wireless Commun.*, 1(4), 660–670, Oct. 2002.
20. O. Younis and S. Fahmy, HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks, *IEEE Transactions on Mobile Computing*, 3(4), 366–379, Dec. 2004.
21. Y. Yao and J. Gehrke, The Cougar approach to in-network query processing in sensor networks, *ACM SIGMOD Record*, 31(3), 9–18, Sept. 2002.
22. S. Chatterjea and P. Havinga, A dynamic data aggregation scheme for wireless sensor networks, in *Proc. Program for Research on Integrated Systems and Circuits*, Veldhoven, the Netherlands, Nov. 2003.
23. V. Mhatre and C. Rosenberg, Design guidelines for wireless sensor networks: Communication, clustering and aggregation, *Elsevier Ad Hoc Networks Journal*, 2(1), 45–63, Jan. 2004.
24. P. Popovski et al., MAC-layer approach for cluster-based aggregation in sensor networks, in *IEEE IWWAN 04*, Oulu, Finland, May 2004.
25. S. Pattem, B. Krishnamachari, and R. Govindan, The impact of spatial correlation on routing with compression in wireless sensor networks, in *ACM/IEEE IPSN 04*, Berkeley, CA, Apr. 2004.
26. S. Nath, P. B. Gibbons, S. Seshan, and Z. R. Anderson, Synopsis diffusion for robust aggregation in sensor networks, in *ACM SenSys 2004*, Baltimore, MD, Nov. 2004.
27. A. Manjhi, S. Nath, and P. B. Gibbons, Tributaries and deltas: Efficient and robust

aggregation in sensor network stream, in *ACM SIGMOD 2005*, Baltimore, MD, June 2005.

28. L. Hu and D. Evans, Secure aggregation for wireless networks, in *Proc. of Workshop on Security and Assurance in Ad hoc Networks*, Orlando, FL, Jan. 28 2003.
29. B. Przydatek, D. Song, and A. Perrig, SIA : Secure information aggregation in sensor networks, in *Proc. of SenSys'03*, pp. 255–265, New York, 2003.
30. H. Çam, S. Ozdemir, P. Nair, and D. Muthuavinashiappan, and H.O. Sanli, Energy-efficient and secure pattern based data aggregation for wireless sensor networks, in *Special Issue of Computer Communications on Sensor Networks*, pp. 446–455, Feb. 2006.
31. W. Du, J. Deng, Y. S. Han, and P. K. Varshney, A witness-based approach for data fusion assurance in wireless sensor networks, in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM '03)*, pp. 1435–1439, San Francisco, 2003.
32. K. Wu, D. Dreef, B. Sun, and Y. Xiao, Secure data aggregation without persistent cryptographic operations in wireless sensor networks, *Ad Hoc Networks*, 5(1), 100–111, 2007.
33. R. Rajagopalan and P.K. Varshney, Data aggregation techniques in sensor networks: A survey, *IEEE Communications Surveys and Tutorials*, 8(4), 4th Quarter 2006.
34. H. O. Sanli, S. Ozdemir, and H. Çam, SRDA: Secure reference-based data aggregation protocol for wireless sensor networks, in *Proc. of IEEE VTC Fall Conference*, Los Angeles, CA, 7, pp. 4650–4654, Sep. 2004.
35. Y. Yang, X. Wang, S. Zhu, and G. Cao, SDAP: A secure hop-by-hop data aggregation protocol for sensor networks, in *Proc. of ACM MOBIHOC'06*, Florence, Italy, May 2006.
36. S. Ozdemir, Secure and reliable data aggregation for wireless sensor networks, *LNCS 4836*, H. Ichikawa et al. (Eds.), pp. 102–109, 2007.
37. D. Westhoff, J. Girao, and M. Acharya, Concealed data aggregation for reverse multicast traffic in sensor networks: Encryption, key distribution and routing adaptation, *IEEE Transactions on Mobile Computing*, 5(10), 1417–1431, October 2006.
38. J. Domingo-Ferrer, A provably secure additive and multiplicative privacy homomorphism, in *Proc. Information Security Conf.*, pp. 471–483, Sao Paulo, Brazil, Oct. 2002.
39. T. Okamoto and S. Uchiyama, A new public-key cryptosystem as secure as factoring, in *Advances in Cryptology—EUROCRYPT'98*, pp. 208–318, Espoo, Finland, 1998.
40. S. Ozdemir, Secure data aggregation in wireless sensor networks via homomorphic encryption, *Journal of The Faculty of Engineering and Architecture of Gazi University*, 23(2), 365–373, 2008.
41. A. Josang and R. Ismail, The beta reputation system, in *Proc. 15th Bled Conf. Electronic Commerce*, Bled, Slovenia, 2002.
42. S. Ganeriwal and M. B. Srivastava, Reputation-based framework for high integrity sensor networks, in *Proc. of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, Washington, DC, pp. 66–77, 2004.
43. S. Ozdemir, Concealed data aggregation in heterogeneous sensor networks using privacy homomorphism, in *Proc. of ICPS'07 : IEEE International Conference on Pervasive Services*, Istanbul, Turkey, pp. 165–168, 2007.
44. C. Castelluccia, E. Mykletun, and G. Tsudik, Efficient aggregation of encrypted data in wireless sensor networks, in *Proc. of Conference on Mobile and Ubiquitous Systems: Networking and Services*, pp. 109–117, Boston, MA, 2005.

第 12 章　无线传感器网络中的分簇技术

在过去十年中，无线传感器网络（WSN）的使用已大大增加，这显示出了在相应的大规模环境下对可扩展和高能效的路由以及数据收集和聚合协议的迫切需求。层次分簇协议（而不是直接的单层通信方案）已经在上述方向被广泛使用。而且，它们可以大大提高整个系统的性能、生命周期和能源效率。在这一章，我们提出在大型的无线传感器网络环境中最先进的相应的层次分簇方法。在 WSN 中分簇的需求是第一动力，给出了一个隐含的层次网络模式的简单描述。并简要探讨基本的优势、目的和设计的挑战。然后介绍一系列适当的分类参数和整体分类方案。在本章的主体部分，简明的呈现目前最重要的无线传感器网络的分簇算法，并根据前面提到的参数和分类方案作出评价。最后，通过一些一般性的评论和在该领域开放性的研究问题总结本章。

12.1　概述

在现今的大多数无线传感器网络应用程序中，整个网络必须要有在恶劣的环境中进行无人操作的能力，在恶劣的环境中，不容易安排或者有效地管理人的访问和检测，有时候甚至是不可行的。根据这个关键的期望，在许多重要的无线传感器网络的传感器节点被采用相对自由的方法随机地部署到感兴趣的地方（例如，用直升机部署），以一个特定的方式形成网络。而且，考虑到整个区域都要被覆盖，传感器的电池能量持续时间短和在部署过程中可能会有损坏的节点，因此需要大量的传感器；因此数以千计的传感器节点将参与是可能的。而且，在这种环境下的传感器它们的资源受限，电池通常不能被再充电。因此，很明显，需要采用专门的能量感知路由和高可扩展性的数据采集协议，以便使网络的生命周期可以在这种环境下保持足够长。

自然地，把传感器节点分组到簇已被研究界广泛采用，在大型的无线传感器网络环境中，它满足了上述可扩展性的目标，实现了高能源利用率，延长了网络的生命周期。对应的分层路由和数据采集协议意味着传感器节点基于分簇的组织，使得数据融合和聚合成为可能，从而节约了大量的能源。在分层网络结构中，每一个簇有一个被称为簇头（Cluster Head，CH）的领导者，通常执行上面提到的特殊任务（融合和聚合），几种常见的传感器节点（Sensor Node，SN）作为成员。

该簇的形成过程最终形成了一个两层的等级结构，CH 节点来自较高一层和簇成员节点来自较低一层。传感器节点周期性地把它们的数据传输到相应的 CH 节

点。该 CH 节点聚合该数据（从而减少了数据包的转发次数），并直接传输或者通过和其他中间 CH 节点的通信把它们传输到基站（Base Stasion，BS）。然而，因为该 CH 节点需要一直将数据发送到比普通节点更远的距离，它们的能源消耗自然更高。为了平衡网络的所有节点的能量消耗，一个常见的解决方案是在每个簇中周期性地选择新的 CH 节点（从而使所有节点随着时间的变换轮流作为 CH 节点）。一个在分簇网络中隐含的分层数据通信的典型例子在图 12-1 进一步说明。

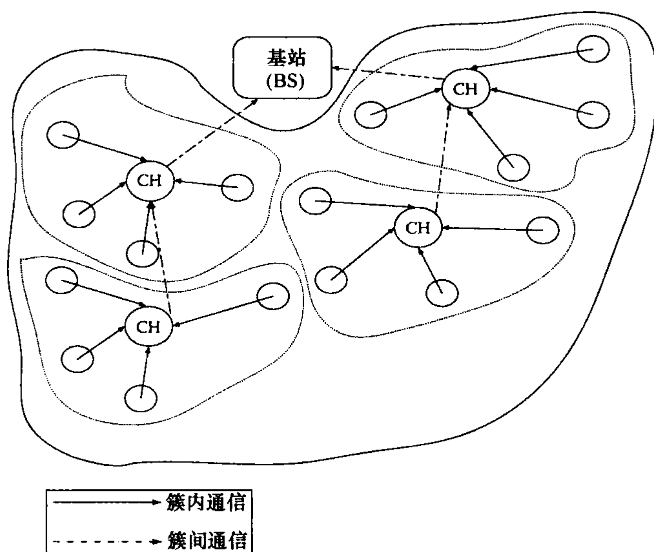


图 12-1　分簇网络中的数据通信

基站是加工来自传感器节点的数据的数据加工点，在基站数据被最终的用户访问。基站通常被认为是固定的，距离传感器节点很远。CH 节点实际上是作为传感器节点和基站间的网关。每一个 CH 节点的功能，正如已经被提到的，是在簇中为所有节点执行常用的功能，像在向基站发送数据前聚合数据。在某种程度上，CH 节点是簇节点的 sink 节点，基站是 CH 节点的 sink 节点。而且，在传感器节点、sink 节点（CH）和基站之间形成的这种结构只要有需要可以重复使用，如果需要的话，分层的无线传感器网络可以创建多层。

12.1.1　无线传感器网络中分簇设计的主要目的和挑战

正如在开始时提到的，WSN 中的层次分簇可以大大提高整个系统的可扩展性、生命周期和能源效率。在簇中分层路由是减少能源开销的一个有效的方法，执行数据聚合和融合是为了减少传输到基站的信息传输量。相反的，一个单层的网络会随着传感器密度的增加而导致网关负载。这样的负载会导致通信的时延和不适当的事件跟踪。而且，因为传感器节点通常不能长距离通信，因此这个单层次的结构对于

那些分布在广阔领域的一大组传感器节点来说是不可扩展的。层次分簇对于那些需要扩展到几百或上千个节点的应用程序来说是非常有用的。在这种情况下的可扩展性意味着负载的均衡和高效的能源利用率。应用程序需要高效的数据聚合（例如在一广阔的地域计算检测辐射的最大值）。路由协议也可以采用分簇[9,27]。在参考文献［50］中，分簇被认为是有效定位目标区域的有用工具。

除了通过数据聚合支持网络的可扩展性和较少能源消耗，分簇还有许多其他的优点和相应的目标[1]。它可以在簇中本地化路由的设置，因此减小了存放在各节点的路由表的大小。同时，因为分簇限制了簇之间相互作用的范围，并避免了节点间多余消息的交换，使得它能够节省通信宽带。而且，分簇能够在传感器级别上稳定网络的拓扑，因此削减了维护网络拓扑的开销。传感器（节点）只关心和它们的CH节点的联系，在层间的CH节点的变化对传感器节点无影响。为了进一步提高网络的运行和延长单个传感器节点的电池寿命以及网络的生命周期，CH节点同样实现了优化管理策略。一个CH能够在簇中安排活动计划以便节点可以切换到低功率的睡眠模式并减少能源消耗率。而且，传感器节点能参与全牌环顺序的调度以及传输和接收时间的确定，从而可以避免传感器的重绑，限制冗余的覆盖，阻止访问冲突。

Wsn同样在设计和实施上避免了几个特殊的问题。在移动自组织的网络（Mobile Ad Hoc Network，MANET）领域，同样面临着相同的挑战和目标，自然地，大量的相关理论（分簇协议等）已经被使用。在WSN中，无论移动支持是可用的还是不重要的，传感器节点的能力受限（电池能源，传输范围，处理硬件和内存使用等）都与特殊的环境相联系（不是简单地为了电池充电或更换所有的传感器就可以访问），这使得能源高效和可扩展性的因素更为重要。而且，根据上述条件，只通过传统的技术，延长网络生命周期的挑战是难以实现的。因此，使用替代的技术（也就是，见12.3节）是不可避免的，从而产生一个更有效的协议，它与MANET的设计有着很大的差异。

除了上述典型的问题（有限的能源、有限的能力、网络的生命周期）在WSN的分簇算法的设计过程中，一些其他的重要的考虑因素如下：①簇信息：CH的选择和簇的形成过程应该产生最佳的簇（很好的平衡，等等）。然而，它们也要保持较低的交换消息的数量，总的时间复杂度应该保持是常数，并独立于网络的增长。这就产生了一个非常具有挑战性的权衡问题。②应用程序的依赖：当为WSN设计分簇和路由协议时，应用程序的健壮性必须得到高度重视，所设计的协议要能够适应各种应用需求。③安全通信：如同传统的网络，数据的安全在WSN中也是同样重要的。在考虑军事应用的网络时，一个WNS分簇方案保持通信安全的能量更为重要。④同步：分时隙的传输方案，例如TDMA允许节点定期地制定睡眠间隔使能源的使用最小化。这种方案需要相应的同步机制，这一机制的有效性必须被考虑。⑤数据聚合：因为这个过程可以使能量优化，它始终是许多传感器网络方案中

一个基本的设计挑战。然而在许多应用程序中它的有效实现并不是一个简单的过程，它必须根据特定的应用环境进一步优化。

12.2　分簇算法分类

12.2.1　分簇参数

在记录 WSN 分簇算法可能的分类选项和算法本身的详细资料前，报告在 WSN 的整个分簇过程中的一些重要参数是有意义的。这些参数同样作为进一步比较和对这一章所提出的分簇协议进行分类的基本方法。

（1）分簇的数量　在最近的概率和随机分簇算法中，CH 节点的选择和形成的过程自然而然导致可变的分簇数量。然而在一些公开的方法中，CH 的设置是预先决定的，因此分簇的数量是预先设定的。对于整个路由协议的效率来说，分簇的数量是一个关键的参数。

（2）簇内的通信　在一些初始的分簇算法中，传感器和它指定的 CH 之间的通信假设是直接的（单跳通信）。然而，簇内的多跳通信现在是经常需要的，也就是说，当传感器的通信范围受限或传感器节点的数量相当大，而 CH 的数量是有限的时，需要进行簇内的多跳通信。

（3）节点和 CH 的移动性　如果我们假设固定的传感器节点和固定的 CH 节点，可以形成一个稳定的分簇，并促进簇内和簇间的网络管理。相反的，如果 CH 或节点是可移动的，每个节点的分簇成员应该动态地改变，迫使分簇随时间而改变，而且可能需要不断地维护。

（4）节点的类型和角色　在一些被提议的网络模型中（也就是说，异构环境），CH 配备了比其他节点更多的计算和通信资源。在许多常见的网络模型中（也就是说，同构环境），所有传感器都具有相同的功能，已部署的传感器中只有一个子集被指定为 CH。

（5）分簇的形成方法　在最近的方法中，当 CH 只是一个普通的传感器节点，时间效率是首要的设计标准时，簇在没有协调的分布式方式下运行。在一些更早的方法中，采用一个集中（或混合）的方法；一个或更多个协调节点用来分割整个网络，控制簇的成员。

（6）簇头的选择　在一些被提议的算法中，簇的领导者节点可以被预先指定。在多数情况下（主要是在同构的环境中），CH 从部署的节点集中选择，选择以概率的方式或者完全随机的方式或基于其他更为具体的标准（剩余能量、连通性等）进行。

（7）算法复杂性　在最近的算法中，执行协议的快速结束是设计的主要目标之一。因此，时间复杂度或大多数分簇的形成过程的收敛速度是一个常数（或只

依赖于CH的数量或跳数的数量)。在一些更早的协议中，时间复杂度允许依赖于网络中传感器的数量，更侧重于其他标准。

(8) 多层次　在一些已发表的方法中，引入了一个多层次分簇等级的概念，用于实现更好的能量分布和整体能量消耗（代替使用单层次的簇)。由多层次分簇实现的促进正在被进一步研究，尤其是当我们有一个非常大的网络和CH内部通信效率很重要时。

(9) 重叠　一些协议也对在不同簇中节点重叠的概念给予了高度的重视（为了更好的路由效率或为了执行更快的分簇形成协议或其他原因)。然而，大部分已知的协议仍然只试图拥有最小的重叠或不支持重叠。

根据上诉参数，接下来我们尝试引入和进一步比较在本章中提出的算法。在表12-1中给出了一个简短的初步介绍。读者应该参考该表并与下一节给出的分类方法相结合，从而获得对提出的算法的更清晰的见解。

表12-1　提出的分簇算法的比较

分簇方法	时间复杂度	节点移动性	簇重叠	簇内拓扑	分簇个数	分簇过程	簇头选择	簇头交替	多层
LBC[5]	N/A	无	无	1-跳	固定	集中式	预设	无	无
MSNDP[6]	N/A	无	无	1-跳	变量	集中式	预设	无	无
LCA[7]	变量	可能	无	1-跳	变量	分布式	基于ID	无	无
AC[9]	变量	有	无	1-跳	变量	分布式	基于ID	无	无
DCATT[10]	N/A	无	无	1-跳	固定	手动式	预设	无	无
LEACH[11]	常量	受限	无	1-跳	变量	分布式	概率/随机	有	无
EEHC[13]	变量	无	无	k-跳	变量	分布式	概率/随机	有	有
HEED[14]	常量	受限	无	1-跳	变量	分布式	概率/能量	有	无
LEACHC[12]	N/A	受限	无	1-跳	变量	集中式	概率/随机	有	无
TLEACH[15]	常量	受限	无	1-跳	变量	分布式	概率/随机	有	有
MOCA[16]	常量	受限	有	k-跳	变量	分布式	概率/随机	有	无
TCCA[17]	变量	无	无	k-跳	变量	分布式	概率/能量	有	无
EECS[18]	常量	无	无	1-跳	常量	分布式	概率/能量	有	无
EEMC[19]	变量	无	无	k-跳	变量	分布式	概率/能量	有	有
RCC[21]	变量	有	无	k-跳	变量	分层式	随机	无	无
CLUBS[22]	变量	可能	有	2-跳	变量	分布式	随机	无	无
FLOC[23]	常量	可能	无	2-跳	变量	分布式	随机	无	无
RECA[24]	常量	无	无	1-跳	变量	分布式	随机	有	无

（续）

分簇方法	时间复杂度	节点移动性	簇重叠	簇内拓扑	分簇个数	分簇过程	簇头选择	簇头交替	多层
HCC[27]	变量	可能	有	k-跳	变量	分布式	连通性	无	有
HC[28]	变量	可能	无	1-跳	变量	分布式	连通性	无	无
MMDC[29]	变量	有	无	k-跳	变量	分布式	连通性	无	无
EEDC[30]	变量	无	无	1-跳	变量	集中式	连通性	无	无
CAWT[31]	常量	无	无	2-跳	变量	分布式	连通性	无	无
EACLE[32]	变量	无	无	2-跳	变量	分布式	邻近	有	无
ACE[33]	常量	可能	有	k-跳	变量	分布式	连通性	无	无
WCA[38]	变量	有	无	1-跳	变量	分布式	基于权重的	无	无
DWEHC[39]	常量	无	无	k-跳	变量	分布式	基于权重的	无	无
TASC[40]	变量	无	无	2-跳	变量	分布式	基于权重的	无	无
GS3[25]	变量	可能	有	k-跳	常量	分布式	预设	无	无
GROUP[26]	变量	无	无	k-跳	受控的	分层式	邻近	无	无

12.2.2 分类簇集协议

已经有一些不同的方法（直接基于上面提到的参数或不是基于上面提到的参数）用于初步区分和进一步对这些用于 WSN 分簇的算法分类[4]。在参考书中，两种最早和常见的分类是：同构或异构网络的分簇算法和集中式或分布式的分簇算法。

上述第一个分类是基于簇中传感器的特性和功能，然而另一个是基于用于形成簇的方法。在异构的传感器网络（例如参考文献，[6，10]）中，有两种类型的传感器，具有更高处理能力和复杂硬件的传感器（普遍用来创建一些在 WSN 内部的骨干——被预置为 CH 节点——同样作为数据采集器和处理由其他节点产生的数据的处理中心）和具有较低能力的普通传感器，它们被用于感知在该领域中感知所需的属性。在同构网络中，所有节点拥有相同的特性、硬件和处理能力（也就是，典型的案例是当传感器节点部署在战场中时）。在这种情况下（在如今的应用中很常见）每个节点都能作为 CH。而且，CH 角色可以在节点间轮转，从而实现更好的负载平衡和更均匀的能量消耗。

同样，当所有节点拥有相同的功能（同构环境）一个分布式的 CH 选择和形成的过程是获得更大的灵活性和更快的执行速度的最合适的技术。还有一些使用集中或混合技术的方法（例如参考文献［5，6，12］，它们具有一个或多个协调节点或基站负责网络的分割和控制簇成员），然而它们对于实际通用的大型 WSN 应用来说是不合适的（它们只适用于特殊用途和有限规模的应用，在应用中需要高质量的连通性和网络分割）。这里我们主要关注同构环境下的分布式的（它对于大型网络更有效）分簇算法（是如今最广泛使用的）。

另一个常见的分类是静态和动态之间的分簇。当一个簇形成过程包括周期性（定期或时间驱动）的CH节点改选或者簇重组过程，用来有效地应对网络拓扑的改变和合适的调整簇集拓扑，或者仅仅着眼于节点间CH角色的恰当轮换来提高能源的效率，那么一个簇的形成过程被认为是动态的。动态的簇结构更好地利用了WSN中的传感器，自然而然改善了能源消耗管理和网络的生命周期。

根据簇形成的条件和CH选择使用的参数，大多数已知的WSN的分簇算法能被进一步分为两类（正如12.3节和12.4节的详细介绍）。

1）概率（随机或混合）分簇算法

2）非概率的分簇算法

在概率选择分簇算法中，先验概率被分配给每一个传感器节点用于确定初始的CH（或者使用一些其他类型的随机选择过程）。分配给每个传感器节点的初始概率作为主要的标准，从而使节点能够单独决定CH（以灵活，统一，快速和完全分布式的方法）；然而，在CH的选择过程中（也就是，剩余能量），或者在簇形成过程中（也就是，通信消耗），其他次要的标准也要考虑，以实现更好的能量消耗和网络生命周期。除了高能效（这从经常采用的周期性改选CH方案中得以促进），这一类别的分簇算法通常实现更快的执行或收敛时间，并减少信息的交换量。

在非概率分簇算法[25-43]中，考虑了更具体的的CH选择和簇形成的标准，它们主要是根据[25-36]节点上的邻近度（连通性，度等）和从附近节点收到的信息。簇的形成过程主要根据节点和它的邻居节点的通信（单跳或多跳邻居节点），通常需要更深入的信息交换，可能在一定程度上遍历地图，因此有时会导致比概率/随机分簇算法更差的时间复杂度。相反地，这些算法对于强劲的方向提取和均衡的分簇更可靠。除了节点邻近度，一些算法[37-40]也使用度量组合，例如剩余能量、传输能量和移动性（形成相应的组合权重），来实现比单一标准协议更普遍的目标。在同一类别中，我们也可以为WSN采用相对较新，颇具挑战性的分簇算法，也就是，生物启发协议[41-43]（基于簇体智能），它可能是如今WSN中最有前景的替代方法。

而且，在12.5小节，我们分别参考了一个特殊用途的分簇协议，它们适合被动型网络[44-49]。这些协议和常见的分簇算法，即所有上述提到的协议相比有明显不同的目标。它们是专门面向具有时间限制的应用，在WSN中通常利用用户对感知数据的查询请求或者在WSN中发生的具体的触发事件。还应当指出在本章的传感器节点或CH移动性的概念在一些特殊的方式中没有被考虑，因为需要移动节点的应用程序数量仍很少。而且在文献中也没有专门的研究。读者可以在参考文献[51，52]找到相关的信息和研究。

最后，在详细介绍上面提到的主要分簇算法前，我们将简要提及在WSN中用于分簇的以前的协议（也就是说，甚至在10年前）。第一个分簇算法的灵感来自于已经被研究，并在无线传感器网络或后来在移动自组织网络领域使用的相应算法。在这些算法中统一分配唯一的标识符是选择CH的关键参数。

第一个这样的分簇算法之一（最初是用于有线传感器网络）是连接分簇算法（Linked Cluster Algorithm，LCA[7]）。LCA 是一个分布式的基于 ID 号、单跳、静态的分簇算法，它试图最大化网络连通性。LCA 的主要缺点是它会导致分簇的数量过多。在参考文献［8］中，提出了一个改善的基于 LCA 的方法（产生较少的分簇数）。因为这两种算法[7,8]都没有考虑 WSN 能源受限的问题，所以作为分簇算法它们具有非常有限的范围。同时，这两个算法构建的是单跳簇，它们的时间复杂度 $O(n)$ 对于大型的 WSN 来说是不可接受的。同样，在参考文献［9］中提出了一个早期的为移动自组织网络开发的，后来用于 WSN 的聚合协议的例子是一种自适应的分簇算法。起初为 MANET 设计的其他一些典型的算法是 MAXMIN[29]，HC[28] 和 WCA[38] 算法；我们将在 12.4 简要地介绍它们。最后，WSN 中最初提出的一些分簇方案是人工分簇的。（主要使用于异构环境中）。这样的典型案例可以在参考文献［10］（DCATT）中找到。这些基于人工形成簇的方法对于现在通用的 WSN 是不可行的，除非满足特殊的条件。

12.3　概率分簇方法

在过去十年中，随着在大区域 WSN 使用中对高效率需求的急剧增加，已经开发了很多具体的分簇协议来满足这个额外的要求（增加网络的生命周期，减少和平衡能源消耗，可扩展性等）。WSN 的分簇协议中最重要和最广泛使用的代表（LEACH、EEHC and HEED），和它们最有价值的扩展将在本章的主要部分呈现。实质上它们都是概率的，它们主要的目标是减少能量的开销，延长网络的生命周期。它们中的一些（例如 LEACH、EEHC 和它们的扩充）在选择 CH 时（初始分配的概率作为 CH 随机选择时的基础）遵循随机的方法，然而其他一些（像 HEED 和相同的方法）遵循混合概率的方法（在 CH 的选择阶段也考虑二级标准，即剩余能量）。一些额外的高效节能的随机选择方法（像 RECA）具有良好的性能，在本章的最后也做了讨论。

12.3.1　广泛的概率分簇协议

12.3.1.1　低能量的自适应分簇层次

其中一个首要的和最流行的分簇协议是 LEACH（低能量自适应分簇层次）[11,12]。它可能是第一个动态分簇协议，它明确提出 WSN 的要求，使用均匀的随机部署的固定传感器节点。并且它是其他改善的分簇协议的基础。它是一个分层的、概率的、分布式的单跳协议，主要目标是①通过平衡网络中所有传感器节点的能量消耗来提高 WSN 的生命周期；②减少节点的能量消耗（通过执行数据聚合来减少传输的消息数量）。它根据收到信号的强度来形成簇，它同样使用 CH 节点作为到基站的路由器。

LEACH通过使用分布式的算法来形成簇，节点不受任何集中控制作出自主决定。所有的节点都有机会成为CH来平衡每个节点的能量消耗。最初，一个的节点拥有概率“p”决定成为CH，并广播它的决定。明确的，在它的选择之后，每个CH广播一条通告消息到其他节点，每个节点（非CH）通过选择使用最少的能量能够到达（基于每个CH消息的剩余能量）的CH来确定属于自己的簇。在图12-2给出了分簇形成过程更明确的说明。

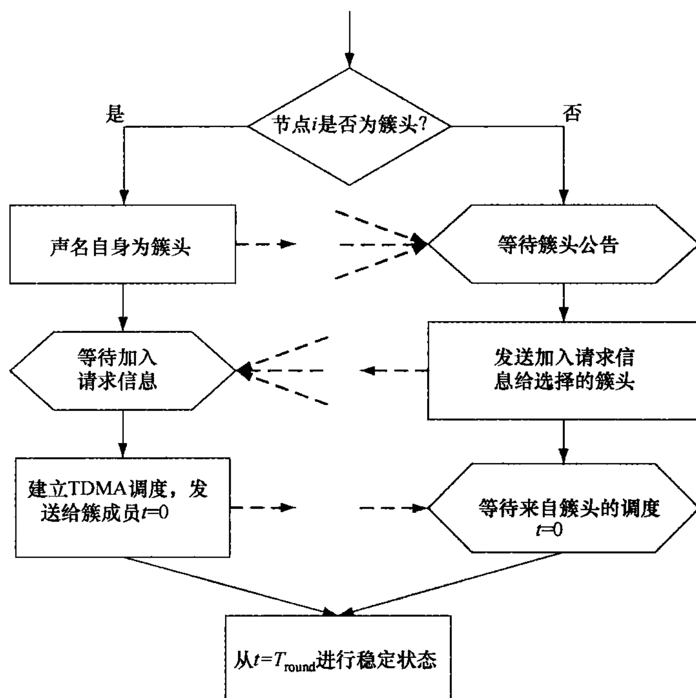


图12-2　LEACH簇集形成过程流程图

CH角色在簇中的传感器节点间定期轮换来平衡负载。这种轮换是通过获得传感器节点选择的一个在0，1之间的随机数“T”来执行的。如果这个数字小于下面的门限，那么节点成为当前轮换的CH节点。

$$T(i)=\begin{cases}\dfrac{p}{1-p\left(r\bmod\dfrac{1}{p}\right)} & i\in G\\[2ex] 0 & \text{其他情况}\end{cases}$$

式中　p——在传感器中所需的CH节点的百分比；

r——本轮的数目；

G——在过去的$\dfrac{1}{p}$轮中还没有成为CH的节点集。

这些簇在每一轮中动态地形成（通过图 12-2 的流程），执行这一轮的时间随机选择。

通常，在单跳网络中 LEACH 能够提供一个相当均匀的负载分配。通过 CH 的随机轮换，它提供了一个良好的平衡的能量消耗。而且，在 LEACH 中使用的本地化的协调方案为簇的形成提供了更好的可扩展性，更好的负载平衡提高了网络的生命周期。然而，尽管 LEACH 普遍表现良好，但是它同样有一些明显的缺点。因为对 CH 的选择和轮换是概率性的，所以一个能量很低的节点也有机会被选为 CH 节点。由于同样的原因，选择的 CH 将集中在网络的一个部分（不能产生良好的 CH 分布），一些节点可能在它们的通信范围之内没有 CH 节点。同样，假设 CH 有很大的通信范围，所以数据可以直接到达基站。但这并不总是一个可以实现的假设，因为 CH 通常是固定的传感器，基站往往不能直接到达所有的节点。而且，LEACH 在一般的单跳簇内和簇间拓扑形成，在这里每一个节点的数据直接传输到 CH，然后传输到基站，因此，通常它不能够在大型区域的网络部署中被有效地使用。

12.3.1.2　节能高效的层次分簇

在参考文献［13］中提出了另一个重要的概率分簇算法［节能高效的层次分簇（Energy-Efficient Hierarchical Clustering，EEHC）］。这个算法的主要目的是通过把簇结构扩展到多跳来克服单跳随机选择网络的缺点，例如 LEACH。它是个分布式的，k 跳层次分簇算法，目标在于实现网络生命周期的最大化。最初，具有概率"p"的节点被选为 CH 节点，向在自己的通信范围内的邻居节点通知它的选择。上述的 CH 现在被称为志愿者 CH。然后，所有距离志愿者 CH 节点 k 跳距离的节点应该直接地或者通过中间节点传输接收到选择的消息。因此，任何收到 CH 选择的消息而不是 CH 的节点就成为最近一个簇的成员。同时，从既不是 CH 又不属于一个簇的节点中选择一些被迫 CH。特别地，当这个选择消息不能在一个预设的时间间隔 t 内到达一个节点，那么，这个节点会成为被迫 CH，它被认为是不属于所有志愿者 CHk 跳内的节点。

然而，EECH 算法最具挑战性的特点是直接扩展到相应的多层次分簇结构。初始的分簇过程在 CH 层次递归地重复，使它能够建立分簇层次的多层次。假设一个 h 层的分簇层次已经使用这种方法构造（每一层相应的预设的 CH 选择概率为 p_1，$p_2,\cdots,p_h$），这个算法保证普通节点和基站之间的高效 h 层通信，如下所示（假设层次 h 是最高的）：普通节点传输它们的收集数据到相应的第一层（层 #1）CH，第一层的 CH 传输它们的聚合数据到第二层的 CH 等，直到到达分簇的最高层；这一层的 CH 传输它们最终聚合的数据到基站。这个多层协议具有 $O(k_1+k_2+\cdots+k_h)$ 的时间复杂度，k_i 是相应的参数（对于每一层）与上面提到的 k 参数（在基本的初始过程中的跳数）相对应。这对于许多现存的算法在那之前（像 LCA）的 $O(n)$ 的时间复杂度来说是一个显著的改进，使得这个算法非常适合于大型网络。

考虑到 EEHC 的整体表现，网络操作的能量开销（数据收集，聚合，传输到

基站等）明显依赖于这个算法的参数 p 和 k 。作者为 p 和 k 的值所推导的表达式实现了最小化能量开销，它们通过仿真结果显示出通过使用最佳参数值，网络中的能量消耗可显著降低。同样仿真结果证实了使用多层（而不是单一层次）的分簇层次的使用价值，正如图12-3所示采用不同的通信半径 r 和空间密度 λ 。

12.3.1.3 混合节能高效的分布式簇集

另一个改进的，非常受欢迎的节能高效的协议是混合节能高效的分布式簇集(Hybrid Energy-Efficient Distributed Clustering，HEED)。HEED是一个分层的、分布式的、分簇方法、每个簇之间的单跳通信模式被保留，然而在CH和基站间允许多跳通信。CH节点根据剩余能量和簇内的通信开销这两个基本的参数来选择。每个节点的剩余能量用于概率选择CH的初始设置。在另一方面，簇内的通信开销反映节点的度或节点到邻居节点的距离，并被节点使用来决定是否加入簇。因此，不像LEACH，在HEED中，CH节点不是随机地被选择的。只有剩余很多能量的传感器节点才能成为CH节点。同样，在彼此的传输半径范围内的两个节点成为CH的概率是小的。不像LEACH，这意味着在网络中CH节点的分布是均匀的。而且，当选择一个簇时，一个节点将和CH节点通信，产生簇内最低的通信开销。在HEED中，每个节点被精确地映射到一个簇，能够直接和它的CH节点通信。同样，所有节点的能量消耗是不均匀的。该算法被分为三个阶段。

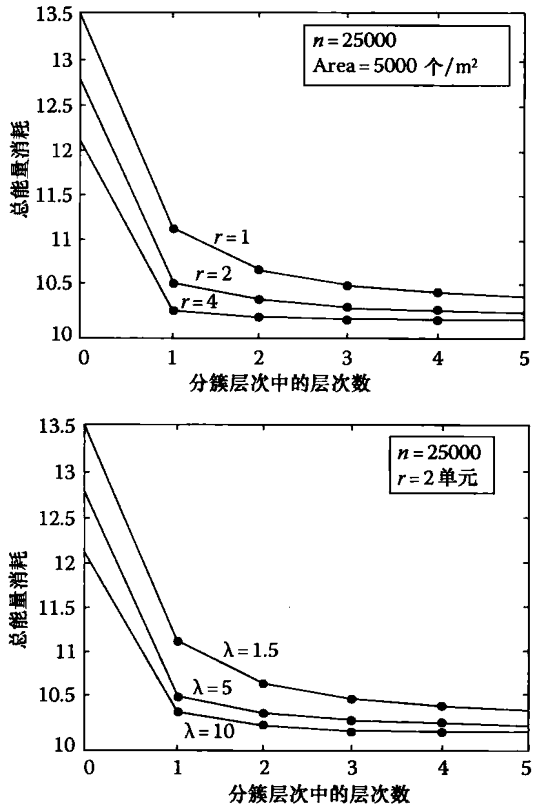


图12-3 多层次EEHC的能量消耗

首先，这个算法在所有节点之间设置一个CH节点的初始百分比。这个百分比的值 C_{prob} 用于限制CH到其他节点的公告。每个节点设置它成为CH的概率，CH_{prob}，如下所示：对 $CH_{prob} = C_{prob} E_{residual}/E_{max}$，这里，$E_{residual}$ 是传感器当前的剩余能量，E_{max} 是最大的能量，它和一个充满电的电池相对应。CH_{prob} 不允许低于某一个值 p_{min}，它被选择为和 E_{max} 成反比。

其次，该算法的主体包括大量（常数）的迭代。每个传感器节点通过这些迭代直到找到CH，可以以最少的通信开销（代价）来传输给它。如果没有CH节点，

这个传感器节点选它自己为 CH 节点，然后向它的邻居节点发送通告告知它们状态的改变。最后，每个传感器把它的 CH_{prob} 值加倍，到达这个阶段的下一个迭代。当 CH_{prob} 达到 1，它停止执行这个阶段。因此传感器节点可以通知它的邻居节点 CH 状态的两种类型：①如果 CH_{prob} 小于 1（如果找到一个更低开销的 CH 在以后的迭代中它可以改变自己的状态）这个传感器成为临时的 CH 节点；②如果 CH_{prob} 的值已经达到 1，这个传感器就成为一个永久的 CH 节点。

最后，每一个传感器节点就其最后的状态作出一个最后的决定。它可以选择低消耗的 CH 或者声明自己作为 CH 节点。另外也要注意，对于一个给定的传感器的通信半径，可以调整 CH 选择的概率以保证 CH 内部节点的连通性。

大体上，HEED 选择簇头节点和形成簇的机制是通过低开销的本地化通信形成一个分布均匀的分簇网络。对于网络的生命周期和所需的能源消耗分布，它明显优于 LEACH。然而，尤其是在大型的网络中，需要同步机制，并且对于距离较远的 CH 传输数据时的能源消耗是非常显著的。同样，需要整个网络的信息来确定可靠的簇内部的通信开销，这些参数的配置在真实世界中可能是不同的。

12.3.2 扩展和其他类似的方法

基于 LEACH 的概率性质，一些其他协议的发展目标在于更好的能源消耗和整体表现。首先，在参考文献 [12] 中提出了 LEACH-C 和 LEACH-F 协议，介绍了稍作修改的初始 LEACH 分簇形成的过程。LEACH-C 是 LEACH 的一个集中式版本，它将分簇创建的责任转移到了基站的身上。每一个节点的初始责任是和基站直接通信以便形成网络的全局视图。因此执行了一个改善的分簇形成过程，实现了稍微好一点的网络整体性能。LEACH-F 同样是一个集中式的协议，它最初是基于相同的全局分簇方案，正如在 LEACH-C 中。主要的区别在于一旦分簇形成，它们是固定的，因此减少了网络中分簇形成的开销。然而，上述设计阻止了任何移动性在网络中的协议的使用。

在参考文献 [15] 中，提出了 LEACH 的一个有价值的扩展（两层 LEACH），概率 CH 选择的关键思想是扩展（和 EEHC 协议相似，但保持簇内单跳拓扑）形成一个两层的分簇方案。外层包括主要的 CH，内层包括次要的 CH 节点。在外层簇的主要 CH 节点和相应的次要 CH 节点直接通信，在内层簇的次要 CH 节点和内层簇的节点直接通信。在一个簇中的数据融合和通信的执行与在 LEACH 中相似，采用 TDMA 方法。主要和次要节点的选择的执行方式和 LEACH 中的一样，为每一个传感器节点设置一个先验概率。首先从剩余节点中选择主要节点，随后选择次要节点。成为主要 CH 节点的概率通常低于成为次要 CH 节点的概率。通常，这个算法的两层分簇方案实现了每轮中向基站传输数据的节点百分比的显著降低。因此，它通常能够减少总的能量消耗。

同样，许多公开的概率分簇算法构建“分离”的簇。相反地，在参考文献

[16] 中作者认为在簇间允许一些重叠度可以很有效地执行一些任务，例如簇间路由，拓扑发现，节点定位，从CH节点故障中恢复等。特别地，他们介绍了一个概率的，分布式多跳重叠的分簇算法（MOCA）用于在重叠分簇中组织传感器。这个分簇过程的目的是保证每一个节点是CH节点或者和一个CH节点的距离在 k 跳之内，k 是预先设置的簇半径。这个算法最初假设每个传感器在网络中以概率 p 成为CH节点。每一个CH向在它的通信范围内的传感器广播自己。这个广播被转发到离CH k 跳之内的所有传感器节点。每个节点向它能够接收到的所有CH节点发送请求加入它们的簇。在加入簇的请求消息中，节点包含所有它能接收到的CH节点的ID号，含蓄地表示它是一个边界节点。CH选择的概率 p 用于控制网络中簇的数量和它们之间重叠的度。作者也提供了大量的模拟工作来验证概率 p 的适当值，从而达到特定的簇计数和重叠度。

除了使用先验概率来选择初始CH节点外，另外一个使用的重要参数是每个节点的剩余能量。在参考文献［17，18］中提出了两个这样的算法（就整体分簇过程而言和LEACH相似)。在参考文献［17］中［时间控制分簇算法（Time Controlled Clustering Algorithm，TCCA)]，整个操作被划分为轮（和LEACH相似)，尝试实现在传感器节点间更好的负载分布。在每一轮发生最初是CH节点选择过程，然后是整个分簇形成的过程。根据剩余能量和预先设定的概率 p 的恰当组合，每个节点决定是否把自己选为CH节点。实际上，在这一步，TCCA通过在每一轮中使部分能量 $E_{residual}/E_{max}$（HEED发起）直接参与计算CH选择最低值 T_i（LEACH发起）将LEACH和HEED直接结合起来。当一个CH被选择，它发送一个消息向它的邻居节点声明它的选择，这条消息包括它的ID号、初始生存时间、剩余能量和时间戳。这个生存时间参数根据剩余能量来选择，它用于限制形成的簇的大小。在另一方面，在参考文献［18］［节能高效分簇算法（Energy Efficient Clustering Scheme，EECS)］中，根据剩余能量和使用本地化的竞争过程无须迭代来完成分簇的形成过程，固定数量的CH节点被选择。特别地，候选的CH通过向邻居候选节点广播它们的剩余能量来竞争成为CH节点的机会。如果一个给定的节点没有找到一个剩余能量更多的节点，那么它就会成为CH节点。同时，通过动态地保持变化的规模，并根据每个簇到基站的距离形成簇。因此，根据离基站很远的簇与离得很近的簇相比需要更多的传输能量这个事实，相应的算法能有效地带来更好的能量消耗和更均匀的负载分布（在仿真试验中和LEACH相比具有更好的表现)。

同样，考虑HEED算法，在参考文献［20］中提出了一个很小的修改。特别地，不同之处在于对最终没有收到任何CH消息的节点的处理；在初始协议的最后阶段，所有的这些节点都成为CH节点。相反地，在参考文献［20］中，作者声称为那些单个的节点重新执行这个算法会有显著的改进。而且，这种轻微的修改结果表明CH数量显著减少，降低了在CH内部通信所需要的路由树的规模，最终产生更快的数据收集过程。

同样的，考虑多层次的 EEHC 算法，在参考文献［19］中提出了一个有价值的扩展（包括额外的 CH 选择标准），在参考文献［19］中每一层预期的 CH 数量由解析公式实现确定。作者概括了在参考文献［13］中提出的分析，并呈现了在一个特定层次最佳 CH 数的结果。考虑到这个形成过程，它们遵循从一级分簇的形成开始的自上而下的方法。每一层的 CH 节点是根据一个特定的概率随机选择。一个节点成为 CH 节点的概率是和节点的剩余能量以及节点到 sink 节点（或者属于最底层的 CH）的距离成正比的。考虑到这个距离，因为每个 CH 节点应该代表它的成员节点发送聚合数据到下一层的 CH 节点，这两个节点间的大的距离将导致在传输 CH 时能量快速消耗。这个概率也是标准化的以便每个层次预期的 CH 的数量等于它们在分析中确定的最优值，因此每个层次预期的 CH 的数量是在它们根据最优值在分析中确定的。也提供了大量扩展的模拟工作，在模拟中 EEMC 协议的结果表明与 LEACH 和 EEHC 协议相比，它能够实现更长的网络生命周期和更短的延迟。

最后，一些随机选择的协议也得到了发展，它们遵循一个更清晰地随机 CH 选择过程（也就是说，通过随机等待或通过产生一个随机的竞争等）。一个早期提出的这种算法是 RCC，它最初是为 MANET 设计的，使用“先声明先获得”的规则。在参考文献［22］中提出了另一种完全随机的分簇算法（CLUBS），每个节点通过从一个固定的整数范围选择一个随机数来参加选择的过程，然后从这个数字默默的倒数。在参考文献［23，24］中提出了最新的另外两个非常有效的（在固定时间内收敛）完全随机的协议。在参考文献［23］中［快速本地分簇协议（Fast Local Clustering Service，FLOC）］，它是一个分布式的协议用最小的重叠产生出相同大小的簇。

另一方面，在参考文献［24］中［环形结构的节能高效的分簇算法（Ring-structured Energy-efficient Clustering Algorithm，RECA）］，在初始分簇形成期间节点被分为单跳架桥簇。首先，一个簇中节点的预期数量是由估计得到的，估计如下：$\gamma = N\pi R^2/A$，N 是网络中节点的总数，A 是网络覆盖的区域，R 是最小的通信范围。一个节点可能选它自己或一个簇成员作为 CH 节点。在每一个时隙，一个节点即不选择它自己作为 CH 节点也不与簇联系，那么它会在 0 ~ 1 之间产生一个随机数，并和 $h = \min(2^r \times 1/\gamma, 1)$ 比较，r 是当前时隙的编号。如果产生的数字小于 h，这个节点成为 CH 节点，并向在它邻域内的节点公布这一消息，否则这个节点等待并接收其他 CH 的公告。在收到 CH 公告后，这个节点把自己和信号最佳的簇联系起来。经过大约 $\lg_2\gamma$ 的时间，在网络中的每个节点就成为 CH 节点或簇成员。RECA 也使用确定的算法在一个簇中轮换 CH 节点。在每一轮，根据簇节点在逻辑环中的位置，它们按次序成为 CH 节点。仿真结果（和 LEACH、HEED 协议相比较）表明 RECA 在每一个簇中的节点之间能够实现能量消耗的均匀分布。就期望的网络生命周期而言，它比 LEACH 和 HEED 表现要好。

最后，由于概率（明显的随机或混合）协议简单，而且能源效率高，它们被认为是无线传感器网络分簇算法的领导者。简单的协议，像 LEACH、EEHC 和

HEED引入了这种低时间复杂度、高能源效率的替代技术，其后许多其他的扩展和类似的概率方法得到发展（大部分是通过扩展或合并上述协议的优点来实现的），它们展现出令人满意的整体性能。就有限的和均衡的能量消耗以及增加网络的生命周期而言，EEMC、EECS、MOCA、TCCA和RECA协议被认为是最新的最具有价值的概率的WSN分簇算法，这些协议的基本缺点（然而在大范围的环境中并不重要）是因为它们的概率本质，所以有时CH节点并不是均匀分布的，CH角色也不是总是统一轮转的，有时还会影响能源消耗的分布。混合概率协议（像HEED和它的扩展）在这方面表现得更好，然而它们（因为需要额外处理过程）和明显的随机协议相比增加了总的时间。

12.4　非概率的分簇方法

用来替代前面章节提出的概率算法，WSN分簇算法的另外一个基本类对于信道之间的竞争、CH的选择和簇的形成主要采用更加确定的标准，这个标准主要根据节点的远近和从其他关系密切的节点接收到的信息。这里簇形成过程主要是基于节点与其邻居节点（单跳或者多跳邻居）之间的通信以及一般需要更加密集的信息交换和可能在一定的程度上遍历图。在以联合加权的形式使用额外的指标（包括剩余的能量、传输功率、移动性等）同样是很有前景的技术，而相比于其他的单指标协议能取得更加广泛的目标。而且，一个更加具有挑战和前景的非概率分簇方法是基于簇体智能地使用，且已经促使构造相应的生物激发的分簇协议，已经证明了这种协议在WSN中能够延长网络的寿命。最后，在很多其他的方法中（在这里没有具体阐述），超出了典型的接近于连通性标准的应用，拟定的协议主要是由特定的基于结构的传感器组织引导的，然后逐步地形成新簇而在正常情况下导致簇具有更多的控制/约束的特性。这样的例子能够在参考文献［25，26］中找到。

12.4.1　邻近节点和基于图的分簇协议

这样一种基于邻近遍历的算法很早就在参考文献［27］中提及［分层控制分簇（Hierarchical Control Clustering，HCC）］。它是一个分布式的多跳分层分簇算法，且同样有效地扩展到形成一个多级别的分簇层次结构。在WSN中任何节点都能够初始化分簇形成过程。算法分两个阶段进行，“树的发现”阶段和“簇的形成”阶段。树的发现阶段基本上是一个以初始节点为根节点的分布式的宽度优先搜索树（Breadth-First-Search，BFS）的形成。每个节点u，在每p个时间单元内广播一个信号，这个信号中装载着它到根节点r的最短的跳数信息。如果通过u的路径是很短的，节点u的邻居节点v将把u当做其父节点，且会更新它到根节点的跳数。广播信号携带父节点ID号、根节点的ID号以及子树的大小。每个节点在其子节点的子树大小改变时更新自身子树的大小。当一个节点的子树达到了尺寸参数k，则开

始了簇形成阶段。节点在其子树上发起簇形成过程。假如子树大小少于 $2k$ ，它对于整个子树将形成一个单一的簇，否则，它将形成多跳簇。簇的大小和重叠的程度同样要考虑。在图 12-4 中，进一步说明了多层次结构。这种方法的时间复杂度是 $O(n)$ ，然而已经证明它能够获得相当平衡的簇，而且也能够很好地控制动态的环境。

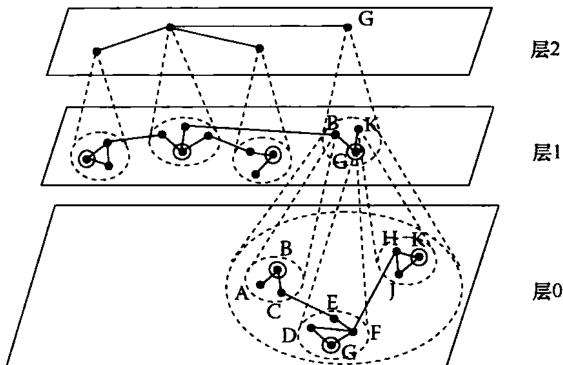


图 12-4　HCC 三层分簇结构

另外两个早期被提到的这种类别的算法（主要为 MANET 设计，然而也在 WSN 中应用）能够在参考文献［28］［高连通性（Highest Connectivity，HC）］和参考文献［29］（最大–最小 D 分簇算法）中找到。在参考文献［28］中提出了一个基于连通性的启发式，这里在邻居集中选择具有最大数量的单跳邻居的节点作为传感器节点的 CH。单跳簇的形成和时钟同步需求限制了这种算法的实际应用。另一方面，在参考文献［29］中，提出了一个分布式算法，这里簇由距离 CH 不超过 d 跳的节点组成。算法具有复杂度 $O(d)$ ，相对于 LCA 和 HC 算法，它不需要时钟同步，并提供一个更好的负荷平衡。

其他基于邻近连通性和邻居信息的算法的更新的例子在参考文献［30-32］中被提及。在参考文献［30］中一个典型的集中式的，基于图的分簇方法（EEDC）被提出。为了最小化簇的数量从而最大限度地节约能量，EEDC 作为一个团覆盖问题模仿簇创建过程，并使用最小数量的团去覆盖图中的所有顶点。sink 节点同样动态地调整基于空间相关性的簇以及从传感器中接收的数据。该算法产生强大和均衡的簇，然而它是集中式的，因此不适合于大规模的 WSN。

在参考文献［31］［等待定时器的分簇算法（Clustering Algorithm via Waiting timer，CAWT）］中，提出了一个基于邻近连通性的分布式算法用于构造具有相同传输距离的同构传感器的层次分簇。一旦传感器被部署，每个传感器向邻居广播一个“hello”信息去表明它的存在，同时侦听其他节点。接收到大量的“hello”信息的传感器（意思是说，节点具有高度连通性）组织成簇，然而其他节点等待形成簇。评价算法的性能使用的是简单的模拟，对于网络寿命具有非常好的表现。然而，很明显，考虑到负荷平衡、CH 重选和在网络中的能量使用，算法的归纳必须经过详细的评估。

相似的，在参考文献［32］中（EACLE）是一个分布式的分簇过程，除了邻近，同样还要考虑到每个节点的剩余能量。这主要是根据两跳邻居的实际传输功率控制方案信息，然后仅仅用簇头建立一个广播树结构。最初，每个传感器都是在等待状态，等待时间 T_1 ，一个在节点的剩余能量上单调递减的函数。当定时器过期，

等待的节点成为一个CH，使用不同的传输功率（高功率和低功率）广播两个包，包里包含广播前接收到的邻居节点的ID。当一个等待节点接收到一个低功率的数据包，它成为一个成员节点，而当其接收到一个高功率的数据包时，它将自己的邻居列表与接收包里的邻居列表的ID进行比较，决定是否是应该继续等待或者变成一个CH。每个节点周期性地执行聚簇过程。一旦节点在一个特定的循环中变成CH，其定时器设置成一个较长的值来避免在下一个周期中又变成CH。

同样，在参考文献［33］［簇建立算法（Algorithm for Clustering Establishment，ACE）］中给出了一个相当有价值的替代方法。与其他的分布式分簇方法不一样，ACE采用了应急算法。应急算法和通过局部优化步骤组合而成的人工神经网络最佳进化方法相类似。最初，节点决定成为CH的候选者，随后广播一个邀请信息。一旦得到邀请，一个邻居传感器加入这个新簇，成为新CH的簇成员。在任何时刻，节点可以属于多个簇。随后，为了选择最佳的候选者成为CH，迁移阶段开始。每个CH周期性地检查邻居是否符合成为CH的性能以及如果在某个邻居节点比其自身有更多的簇成员的情况下而决定不做CH。拥有最大数量的簇成员和与其他簇具有最少的重叠的节点被认为是CH的最佳的最终候选者。算法在时间$O(d)$内收敛，d是单位圆内的节点密度。ACE的试验验证表明当与像参考文献［7，8］中基于节点ID号方案相比较，其获得更低的方差和高的分簇平均大小。

最后，向着高效的数据收集和聚合方向，一些其他的替代方案，在没有直接分簇的情况下，同样被提出来[34,35]。这些方法主要是基于图遍历的启发式方案，并且与传统的基于分簇的执行，例如LEACH相比，显示出相当有价值的改进。同样的，在参考文献［36］中作者提出了一个相对应“混合”分簇协议（PEACH），它在没有给分簇形成过程中带来像其他直接分簇算法中普遍面临的额外的开销的情况下建立一个自适应的分簇层次。相反，通过监听邻居节点传输和接收到的数据，每个节点都能够确定自身的角色（CH或者簇成员），并且可能加入其他节点的簇，因此创造一个分簇层次。在PEACH协议中，当节点接收到以节点本身为地址的数据包，它变成CH。另外，当数据包的目的地是另外一个节点，那么这个监听数据包的节点将会加入那个目的地节点的分簇。通过模拟手段，PEACH协议与其他竞争方法的协议相比较，例如LEACH、HEED和PEGASIS，并表现出了更加低的能量消耗和更高的网络寿命，这主要是因为当前分簇层次的建立是基于节点监听到的信息。

总之，节点邻近和图遍历分簇协议获得了相当平衡和稳定的分簇（在整个区域内CH相当均匀地分布，簇内具有低的通信消耗等），然而它们也存在很多的缺点。它们通常会增加时间复杂度以满足更高的定性标准，然而重要的参数，例如分簇的数量和每个簇的大小在没有质量成本情况下是不容易控制的。另外，相对于概率协议大多数这种协议没有运用有效的CH转换程序（或者一点都没有应用这个程序），而使得能量效率降低以及网络寿命缩短。HCC、ACE、CAWT和EACLE协议

可以被认为是在这种类型中最有价值的方法，然而它们并不和相应的概率簇协议一样，表现不是很好（在能量消耗和网络寿命方面）。相反地，对于当前混合分簇协议需要引起特别的注意（没有直接的分簇），例如 PEACH 协议，它们似乎能够通过有效地避免一部分簇形成开销来获得更低的能量消耗。

12.4.2 基于权的簇协议

除了邻近节点，一些其他已知的算法使用一些指标的组合，例如剩余能量，传输功率（因此形成相应的组合权重）去获得与单标准协议相比更广泛的目标。很多遵守这个指令的算法最初借鉴于移动 ad hoc 网络场景，如参考文献 [37, 38]。作为一个典型的例子，在参考文献 [38] 中（WCA），提到了一个相应的基于权重的协议中，这里 CH 选举过程是基于每个节点的“总权重”的计算，它考虑到了很多系统参数，例如节点度、传输功率、移动性和节点的剩余能量：$W_v = w1T_v + w2D_v + w3M_v + w4P_v$。计算总的权重以及由每个节点进行传播。在其邻居中具有最低权重的节点被选为 CH。CH 竞争是一个非周期性的过程；它每次根据需求调用网络拓扑重置是无可避免的。这个算法通过在簇中减少传感器的数量试着去提供更好的负荷平衡，但是时钟同步的需求限制了它的应用。

最近两个基于权重的协议在参考文献 [39, 40] 中提及。在参考文献 [39] [分布式的基于权重的能量有效的层次分簇算法（Distributed Weigh-Base Energy-Efficient Hierarchical Clustering, DWEHC）] 中，提出了一个相应的分布式算法，主要目标是通过形成平衡的分簇规模以及优化簇内的拓扑结构来达到高能源效率。每个节点在自身区域定位邻居之后计算它的权重。权重是传感器剩余能量和到邻居距离的函数。在邻居中，有最大权重的节点被选择为 CH，剩余节点变成一般成员。在这个阶段节点被认为是第一级成员，因为它们和 CH 直接相连。节点使用最少的能量逐渐调整这种成员来到达 CH。基本上，节点检查它的非 CH 邻居去寻求它们到达 CH 的最低成本。给定节点到其邻居的距离，它能够评估是保持在第一级成员更好还是通过两跳路径变成第二级成员达到 CH 更好。图 12-5 说明了簇内拓扑的结构。与 HEED 相比，已经表明 DWEHC 算法能够形成更加均匀的分簇平衡以及获得在簇内和簇间通信中获得相当低的能量消耗。

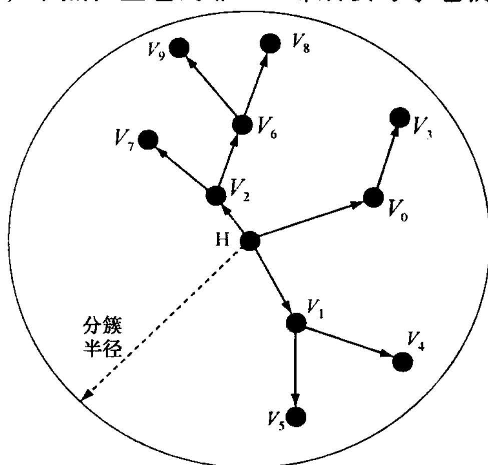


图 12-5　DWEHC 多跳簇内技术

相似地，在参考文献 [40] 中 [拓扑自适应空间分簇（Topology Adaptive Spatial Clustering, TASC）]，作者提出了另外一种分布式算法把网络

分为一些本地各向同性，不重叠的簇而对于簇的数量、簇大小以及节点坐标没有预见性的了解。这通过派生一系列权重包括距离、连通性以及每个节点位置的密度信息获得。派生权重形成了持有一个协调领导者选举程序的地形，从而每个节点选择靠近中心的节点作为其领导者。

总体上，类似于邻近节点和基于图的协议，基于权重的分簇协议已经被证明以一种更加系统和确定的方式形成均衡和稳定的分簇。另外它们能够达到更好的能源消耗分布，因为在选择 CH 的过程中，它们中的大多数将剩余能量用于部分权重计算。然而，它们通常含有与在上一段所说的节点邻近和基于图协议一样的缺点（更长的通信时间，没有 CH 转换等）。

12.4.3　生物激活分簇方法

最后，近年来，也提出了一些新的基于群体智能技术的算法，这种技术模仿社交昆虫，例如蚂蚁的群体性行为。关于网络的生命周期它们已经达到了非常好的仿真实验结果。在参考文献［41］中作者基于 ANTCLUST 方法提出一个基于群体智能分簇算法。ANTCLUST 是一个蚂蚁群体模型，解决这类分簇问题。在群体者模型中，当两个对象集合在一起通过互换和比较它们之间的信息使它们知道彼此是否属于相同的组。在 WSN 情况下，最初含有更多剩余能量的传感器节点独立成为 CH。接着，随机选择的节点相互碰到，交换信息，然后簇建立，合并，并通过这些本地的会合和比较它们的丢弃信息。含有较少剩余能量的每个节点根据特定标准选择一个簇，类似于 CH 的剩余能量，到 CH 的距离，以及估计的簇的大小。最终，形成能量高效的簇，延长 WSN 的寿命。

另一个保证 CH 具有较好分布和高能效的相关方法能在参考文献［42］中找到。同样，在参考文献［43］中，一个协议被呈现并通过仿真进行评价，它的目的是使用粒子群优化（Particle Swarm Optimization，PSO）来最小化簇内的距离和优化网络的能量消耗。一般来说，生物激活分簇算法表明它们能够动态地控制 CH 选择，同时能够获得相当均匀的 CH 分布和能量消耗。然而，就像文章中指出的一样，它们必须进行进一步的研究。

12.5　反应网络的分簇算法

在前面的章节中（被视为“主动的”分簇算法）提到的所有的算法是基于假设传感器一直有数据需要发送的，因此，它们在分簇的形成过程中都应考虑。相反地，被动式算法利用发生在网络中的对感知数据或者指定的触发事件的用户请求。换句话说，节点可能即时响应感知属性突然的或者显著的变化。这个方法对于实时应用是有益的，但不是特别适合于那些需要定期进行数据检索的应用。

阈值感知能量有效的传感器网络协议（Threshold sensitive Energy Efficient sensor

Network protocol, TEEN)[44]形成了一个层次的分簇结构，邻近的节点组合成簇。协议关注于信息聚合而不是簇的形成，这和 LEACH 是非常相似的。协议定义了两个阈值：对感知属性的阈值（绝对的）是硬阈值，而软阈值是属性感知阈值（微小变化）。阈值的概念在 WSN 的应用中是非常重要的，例如火警和温度监控。仅仅当传感器读数达到硬阈值之上和一些定量的变化（软阈值）的情况下，节点会发送这个读数。在 TEEN 中，传感器节点持续不断地感知媒介，但是数据采用低频率传输有利于能量的节约。然而，假如没有超过阈值，节点将不再通信，换句话说 TEEN 将不支持定期报告数据。

自适应周期性的 TEEN（Adaptive Periodic-TEEN, APTEEN）是 TEEN 的变种，它避免了后者的主要缺点。它是一个混合的路由协议，在其中节点仍然对时间紧急情况响应，而且在定期间隔中以能源有效的方式给出网络的一个整体描述。在 APTEEN 中的 CH 选择是基于 LEACH-C 中使用的机制。簇的有效期为一个区间叫做簇周期。在这个周期的末尾，BS 执行重新建立簇。假设相邻的节点存储相似的数据，APTEEN 形成节点对，仅仅是其中的一个节点对查询进行响应。这两个节点能够在查询控制作用中交替，因此节约能量。仿真结构已经表明 APTEEN 的性能在平均能耗和网络寿命方面居于 TEEN 和 LEACH 协议之间。这是来自于 APTEEN 混合前摄的反应特性的一个合理的结论。TEEN 和 APTEEN 的主要缺点是与多层次分簇形成相联系的控制开销，执行的方法是阈值函数，并且它们没有利用空间和时间数据的相关性去提高效率。

分布式的反应分簇算法（Decentralized Reactive Clustering, DRC）在参考文献[48]中提及。与其他的反应算法类似，仅当事件检测时才会发起分簇过程。定义了四个不同的操作阶段：快速部署阶段，其次是簇形成阶段它发生在簇建立好以后，然后是一个分簇内的数据处理阶段以及最后的 CH 处理过程中心阶段。DRC 在分簇形成过程中使用功率控制技术来最小化能量消耗。不幸的是，模拟仅仅把 DRC 与 LEACH 作对比，因此没有突出其与其他的被动簇协议相比的性能提升或者缺点。

最近，提出了分簇聚集（Clustered AGgregation, CAG）[46]理论，它利用传感器数据的空间相关性进一步减少传输数据量，为聚集查询提供近似的结果。CAG 保证结果在用户指定的容错阈值内。当查询分发到网络（查询阶段）中时执行簇形成，簇群节点感知到相似的值。随后，CAG 进入反应阶段在那里每个簇仅仅将一个聚合值传输到聚合树上。事实上，CAG 是个有损耗的分簇算法（大多数感应读数从来没有被报告过），在显著的能量、存储器、计算以及节省通信方面会得到更低的准确度。

改进后的 CAG 算法[47]扩展了 CAG，根据动态的环境定义了两个操作模式。在交互式的模式中，用户发出一个查询后网络形成一个单一的响应。这适合实际情况是动态变化的环境，以及当用户希望交互式地去改变近似粒度或者查询属性。另一方面，在流模式下，CH 对查询只发送一次响应流。这种操作模式非常适合于传感器读数不会频繁改变以及在一个特定的时间周期内查询仍然有效的静态环境。注意交互式模式仅仅利用传感器数据的空间相关性去形成簇，而流模式则充分利用时间

和空间相关性。后者的簇调整随着本地的数据和拓扑变化而进行调整。总体上，CAG和改进的CAG方法提供了有效的数据聚集和节约能量方案，这是以用户提供的精确的错误限制阈值为代价的。

在更近的方法中，Guo和Li提出了一个动态分簇响应路由算法（Dynamic-Clustering Reactive Routing，DCRR）[49]。DCRR借用了生物神经网络的想法，以下的观察表明后者同样适用了多对一（神经元到大脑）的通信模式，这和WSN中的节点相似。在DCRR中，一旦出现事故，根据节点的剩余能量在事件区域中动态地选择CH。对于发送数据到sink节点，DCRR定义了一个TEEN发起的“行动阈值”。阈值是动态地调整去跟踪事件速度的变化。行动阈值同样在事件区域外有意地波动从而使得整个网络节点能够定期地发送数据（因此它们不会被误解为失败的）。

总体上，反应分簇算法减少了WSN中的控制开销和执行分簇过程需要的时间，从而降低了网络开销。使用阈值属性以及开发许多网络环境中固有的感知数据的时空相关性压缩了事件驱动数据传输的幅度。这些算法同样对于剧烈的变化有快速的响应，这对时间敏感应用相当的有益。然而，它们不适合需要定期检索感知读数的应用，这些应用中建立稳定和能效高的簇结构是至关重要的。

12.6　结论

一般来说，WSN中的分簇在过去几十年里引起了极大的兴趣以及有大量的已经发表的相关研究。这章中我们试着去阐明大多数至今在文献中仍然提及的最重要的协议的主要特性。正如有人指出，划分节点到簇中，导致了分层路由和数据聚集协议，被认为是在WSN中支持可扩展性最有效的方法。分层路由结构有助于独立于WSN增长的高效的数据聚集和聚合，总体上降低了总的通信量以及能量消耗。

大多数现存协议的主要目标是基于怎样延长网络的寿命以及怎样对关键资源更加有效地应用，例如电池能量。而且，对于快速收敛时间和最小能耗（考虑簇的形成过程）的结合需求导致了适当的快速分布概率（明确随机或者混合），分簇算法很快就在这个领域内变得流行以及广泛使用。在这些算法中假设节点做出快速的抉择（即是否变成CH）仅仅是根据一些概率或者其他本地的信息（即对于它们的剩余能量），且通常最终簇输出的质量仅被认为是次要的参数。大多数算法（能量消耗得更加均匀地分布）的另外一个特性是在网络中所有节点定期重选CH。分簇算法采用其他经典参数作为最初竞争标准，例如连通性、节点的邻近度、距离等，而且已经得到发展，相关的协议也仍在使用中，可能会导致更加定性的输出（更加均衡的分簇等）。然而因为引用了概率/随机的分簇算法，这些协议的时间复杂性很难保持在较低的水平。

而且，因为在实际应用中的WSN的规模（传感器的数量、覆盖区域等）变得越来越大，多跳通信方式（关于簇内或者簇间通信）的扩展是不可避免的，而多

级别的簇层次分类同样被认为是独立于网络规模的发展而维持能量效率的一个好的选择。对于时间紧急应用的专门协议已经取得了显著的进展，在这些应用中节点应该在传感器属性值方面对突发情况和剧烈的改变做出即时反应。

最后，一些其他的问题在未来的研究中应该进一步探讨。一些具有挑战性的问题包括为了最大限度地提高能量效率而开发一个寻找最优化的分簇数量的一般方法，为了获得更好的能量分布，估计 CH 转换/重选的最优化频率，同时还有，保持较低的总的开销，支持有效的节点和 CH 移动性以及支持移动的 sink 节点，包含一些安全的方面（即，当在恶劣的环境下使用基于分簇的协议加强安全保护的需要），在 CH 失败的情况下进一步发展有效的恢复协议等。

参考文献

1. A.A. Abbasi and M. Younis, A survey on clustering algorithms for wireless sensor networks, *Computer Communications*, 30, 2826–2841, 2007.
2. K. Sohrabi et al., Protocols for self-organization of a wireless sensor network, *IEEE Personal Communications*, 7(5), 16–27, 2000.
3. R. Min et al., Low power wireless sensor networks, in *Proceedings of International Conference on VLSI Design*, Bangalore, India, January 2001.
4. M. Liliana, C. Arboleda, and N. Nidal, Comparison of clustering algorithms and protocols for wireless sensor networks, in *Proceedings of IEEE CCECE/CCGEI Conference*, Ottawa, Ontario, Canada, pp. 1787–1792, May 2006.
5. G. Gupta and M. Younis, Load-balanced clustering in wireless sensor networks, in *Proceedings of the International Conference on Communication (ICC 2003)*, Anchorage, AK, May 2003.
6. E.I. Oyman and C. Ersoy, Multiple sink network design problem in large scale wireless sensor networks, in *Proceedings of the IEEE International Conference on Communications (ICC 2004)*, Paris, June 2004.
7. D.J. Baker and A. Ephremides, The architectural organization of a mobile radio network via a distributed algorithm, *IEEE Transactions on Communications*, 29(11), 1694–1701, 1981.
8. A. Ephremides, J.E. Wieselthier, and D.J. Baker, A design concept for reliable mobile radio networks with frequency hopping signalling, *Proceedings of IEEE*, 75(1), 56–73, 1987.
9. C.R. Lin and M. Gerla, Adaptive clustering for mobile wireless networks, *IEEE Journal on Selected Areas Communications*, 15(7), 1265–1275, 1997.
10. W.P. Chen, J.C. Hou, and L. Sha, Dynamic clustering for acoustic target tracking in wireless sensor networks, in *Proceedings of the 11th IEEE International Conference on Network Protocols (ICNP'03)*, pp. 284–294, November 4–7, Atlanta, GA, 2003.
11. W.R. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, Energy efficient communication protocol for wireless microsensor networks, in *Proceedings of the 33rd Hawaiian International Conference on System Sciences (HICSS-33)*, pp. 3005–3014, January 4–7, Maui, HI, 2000.
12. W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, An application specific protocol architecture for wireless microsensor networks, *IEEE Transactions on Wire-*

less Communications, 1(4), 660–670, 2002.

13. S. Bandyopadhyay and E. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in *22nd Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM 2003)*, San Francisco, CA, April 2003.
14. O. Younis and S. Fahmy, HEED: A hybrid, energy-efficient, distributed clustering approach for Ad Hoc sensor networks, *IEEE Transactions on Mobile Computing*, 3(4), 366–379, 2004.
15. V. Loscri, G. Morabito, and S. Marano, A two-level hierarchy for low-energy adaptive clustering hierarchy, in *Proceedings of IEEE VTC Conference 2005*, Vol. 3, pp. 1809–1813, 2005.
16. A. Youssef, M. Younis, M. Youssef, and A. Agrawala, Distributed formation of overlapping multi-hop clusters in wireless sensor networks, in *Proceedings of the 49th Annual IEEE Global Communication Conference (Globecom06)*, San Francisco, CA, November 2006.
17. S. Selvakennedy and S. Sinnappan, An adaptive data dissemination strategy for wireless sensor networks, *International Journal of Distributed Sensor Networks*, 3(1), 23–40, 2007.
18. M. Ye, C. Li, G. Chen, and J. Wu, EECS: An energy efficient clustering scheme in wireless sensor networks, in *Proceedings of IEEE International Performance Computing and Communications Conference (IPCCC'05)*, pp. 535–540, April 7–9, Phoenix, AZ, 2005.
19. Y. Jin, L. Wąng, Y. Kim, and X. Yang, EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks, *Computer Networks Journal*, 52, 542–562, 2008.
20. H. Huang and J. Wu, A probabilistic clustering algorithm in wireless sensor networks, in *Proceedings of IEEE 62nd VTC Conference*, Dallas, TX, September 2005.
21. K. Xu and M. Gerla, A heterogeneous routing protocol based on a new stable clustering scheme, in *Proceeding of IEEE Military Communications Conference (MILCOM 2002)*, Anaheim, CA, October 2002.
22. R. Nagpal and D. Coore, An algorithm for group formation in an amorphous computer, in *Proceedings of the 10th International Conference on Parallel and Distributed Systems (PDCS98)*, Las Vegas, NV, October 1998.
23. M. Demirbas, A. Arora, and V. Mittal, FLOC: A fast local clustering service for wireless sensor networks, in *Proc. of Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS04)*, Florence, Italy, June 2004.
24. G. Li and T. Znati, RECA: A ring-structured energy-efficient clustering architecture for robust communication in wireless sensor networks, *International Journal Sensor Networks*, 2(1/2), 34–43, 2007.
25. H. Zhang and A. Arora, GS3: Scalable self-configuration and self-healing in wireless networks, in *Proceedings of the 21st ACM Symposium on Principles of Distributed Computing (PODC 2002)*, Monterey, CA, July 2002.
26. L. Yu, N. Wang, W. Zhang, and C. Zheng, GROUP: A grid-clustering routing protocol for wireless sensor networks, in *Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM'06)*, pp. 1–5, September 22–24, Wuhan City, China, 2006.

27. S. Banerjee and S. Khuller, A clustering scheme for hierarchical control in multi-hop wireless networks, in *Proceedings of the 20th Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01)*, pp. 1028–1037, April 22–26, Anchorage, AK, 2001.
28. A.K. Parekh, Selecting routers in ad-hoc wireless networks, in *Proceedings of SBT/IEEE International Telecommunications Symposium (ITS'94)*, pp. 420–424, August 22–25, Rio de Janeiro, Brazil, 1994.
29. A. Amis, R. Prakash, T. Vuong, and D. Huynh, Max-min D-cluster formation in wireless Ad Hoc networks, in *Proceedings of the 19th Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'00)*, pp. 32–41, March 26–30, Tel-Aviv, Israel, 2000.
30. C. Liu, K. Wu, and J. Pei, A dynamic clustering and scheduling approach to energy saving in data collection from wireless sensor networks, in *Proceedings of the 2nd Annual IEEE Conference on Sensor and Ad Hoc Communications and Networks (SECON'05)*, pp. 374–385, September 26–29, Santa Clara, CA, 2005.
31. C. Wen and W. Sethares, Automatic decentralized clustering for WSNs, *EURASIP Journal on Wireless Communications and Networking*, 5(5), 686–697, 2005.
32. K. Yanagihara, J. Taketsugu, K. Fukui, S. Fukunaga, S. Hara, and K.I. Kitayama, EACLE: Energy-aware clustering scheme with transmission power control for sensor networks, *Wireless Personal Communications*, 40(3), 401–415, 2007.
33. H. Chan and A. Perrig, ACE: An emergent algorithm for highly uniform cluster formation, in *Proceedings of the 1st European Workshop on Sensor Networks (EWSN)*, Berlin, Germany, January 2004.
34. S. Lindsey and C.S. Raghavendra, PEGASIS: Power-efficient gathering in sensor information networks, Computer Systems Research Department, the Aerospace Corporation, Vol. 3, pp. 3-1125–3-1130, Los Angeles, CA, 2002.
35. H.O. Tan and I. Korpeoglou, Power efficient data gathering and aggregation in wireless sensor networks, *Issue on Sensor Networks Technology, ACM SIGMOD Record*, 32(4), 66–71, 2003.
36. S. Yi, J. Heo, Y. Cho, and J. Hong, PEACH: Power-efficient and adaptive clustering hierarchy protocol for WSNs, *Computer Networks*, 30, 2842–2852, 2007.
37. S. Basagni, Distributed clustering for ad hoc networks, in *Proceedings of International Symposium on Parallel Architectures, Algorithms & Networks (ISPAN'99)*, pp. 310–315, June 23–25, Fremantle, Australia, 1999.
38. M. Chatterjee, S.K. Das, and D. Turgut, WCA: A weighted clustering algorithm for mobile ad hoc networks, *Clustering Computing*, 5, 193–204, 2002.
39. P. Ding, J. Holliday, and A. Celik, Distributed energy efficient hierarchical clustering for wireless sensor networks, in *Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems(DCOSS05)*, Marina Del Rey, CA, June 2005.
40. R. Virrankoski and A. Savvides, TASC: Topology adaptive clustering for sensor networks, in *Proceedings of the Second IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, (MASS 2005)*, Washington, DC, November 2005.
41. J. Kamimura, N. Wakamiya, and M. Murata, A distributed clustering method for energy-efficient data gathering in sensor networks, *International Journal on Wireless and Mobile Computing*, 1(2), 113–120, 2006.

42. S. Selvakennedy, S. Sinnappan, and Y. Shang, A biologically inspired clustering protocol for wireless sensor networks, *Computer Communications* 30, 2786–2801, 2007.
43. N.M. Abdul Latiff, C.C. Tsimenidis, and B.S. Sharif, Energy-aware clustering for wireless sensor networks using particle swarm optimization, in *IEEE Intl. Symposium PIMRC'07*, pp. 1–5, Athens, Greece, September 2007.
44. A. Manjeshwar and D.P. Agrawal, TEEN: A routing protocol for enhanced efficiency in wireless sensor networks, in *Proceedings of the 15th International Parallel and Distributed Processing Symposium*, San Francisco, CA, 2001.
45. A. Manjeshwar and D.P. Agrawal, APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks, in *Proceedings of International Parallel and Distributed Processing Symposium (IPDPS'02)*, pp. 195–202, April 15–19, Fort Lauderdale, FL, 2002.
46. S. Yoon and C. Shahabi, Exploiting spatial correlation towards an energy efficient Clustered AGgregation technique (CAG), in *Proceedings of IEEE International Conference on Communications (ICC'05)*, pp. 82–98, May 16–20, Seoul, Korea, 2005.
47. S. Yoon and C. Shahabi, The Clustered AGgregation (CAG) technique leveraging spatial and temporal correlations in wireless sensor networks, *ACM Transactions on Sensor Networks*, 3(1), Article #3, March 2007.
48. Y. Xu and H. Qi, Decentralized reactive clustering for collaborative processing in sensor networks, in *Proceedings of the 10th International Conference on Parallel and Distributed Systems (ICPADS'04)*, pp. 54–61, July 7–9, Newport Beach, CA, 2004.
49. B. Guo and Z. Li, A dynamic-clustering reactive routing algorithm for wireless sensor networks, *Wireless Networks Journal*, Springer, 15(4), 423–430, 2009.
50. D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, Next century challenges: scalable coordination in sensor networks, in *Proceedings of the ACM/IEEE MOBICOM Intl. Conference*, Boston, MA, pp. 6–11, August 2000.
51. C.M. Liu and C.H. Lee, Power efficient communication protocols for data gathering on mobile sensor networks, in *IEEE 60th VTC*, Los Angeles, CA, pp. 4635–4639, 2004.
52. X. Zhang, H. Wang, and A. Khokhar, An energy-efficient data collection protocol for mobile sensor networks, in *IEEE 64th VTC*, Montreal, Quebec, Canada, pp. 1–5, 2006.

第13章　无线传感器网络中能量有效的感知行为

在无线传感器网络（WSN）中，由于感知节点主要依赖于电池供电，能量有效性成为一个主要考虑的问题。一个感知节点通常由两部分组成：负责测量的感知单元和进行无线通信的无线通信单元。尽管已经提出了很多无线通信单元的节能模式，但是对于能量感知单元的节能模式很少被提及。最近的测量结果表明感知节点进行感知数据测量时会消耗大量的能量，并且感知的能量消耗和数据通信的能量消耗是相当的。因此，研究感知单元的节能模式是必需的。

无线传感器网络部署应用于监测静态或者动态事件。静态事件（例如温度和湿度）很容易获取。与之相反，动态事件是相当地不合作的，因此很难测量。在战场上，敌方车辆的移动和海洋中鲸鱼的迁徙都是不合作事件的例子。由于它们频繁移动，很难被监测到。因此，只能通过传感器对周围环境的持续监测才能够对它们进行观察。然而，如果感知设备经常处于工作状态，能量有效性成为将要面临的一个严肃的问题。这一章回顾现有的无线传感器网络中的各种感知模式，并研究了能量有效性和事件覆盖的权衡。我们以单一传感器结点以固定的感知休眠时间模型开始，就事件检测失败率和正常平均能量消耗进行测量。如果事件的统计数据是未知的，一个基于倍增/倍减（additive increase/multiplicative decrease，AIDM）规则的自适应模式将被提出来适应感知数据的睡眠调度。将固定模式和自适应模式扩展到多传感器结点，同时除了考虑到能量有效性也会考虑到整个网络的覆盖范围。本章还讨论了感知单元与无线通信单元之间的交互。实现了不相关事件的模型来证明所提出模式的有效性。

13.1　概述

无线传感器网络（WSN）是新兴的无线通信领域广泛应用的新技术之一，例如环境监测、健康监护、智能建筑以及战场控制[1]。一个典型的WSN由能源受限的传感器节点组成，它们负责监测物理现象，并将其报告给访问点或者融合中心。WSN的一个主要特征是长期有效性。因为可能需要将大量传感器节点部署到大面积的多变的地理区域中，电池的更换是不大可能的。WSN必须利用最少的能量在大范围的操作场景中运作。因此，网络寿命成为设计WSN的一个关键问题。

尽管先前的许多工作提到了能量有效的无线传感器网络通信模式（媒体访问控制和路由模式），能量有效感知并没有受到重视。然而，在一个节点中感知单元

消耗的能量与无线通信单元消耗的能量是相当的。再者，感知频率通常比无线通信频率要高。因此，本章将会回顾和研究能量有效的感知技术。首先，在13.2节中我们将会回顾各种节能模式，尤其是能量有效感知。然后，接下来的章节将会对能量有效感知进行系统地分析。在13.3节中，将会研究单个传感器节点的交替感知模式，同时也会考虑固定和自适应模式。两种模式的模拟结果在13.4节中提出。在13.5节中将会介绍完全的网络覆盖。13.6节将会讨论开放性的问题和争议。13.7节总结本章，同时提出对将来这一领域研究工作的建议。

13.2 节能模式回顾

为了达到较长的网络寿命，提出了许多WSN的节能模式。本节对其中三类方法进行了回顾，即硬件能量管理、能量有效的感知技术以及能量有效的无线通信技术。这些技术的例子，如图13-1所示。

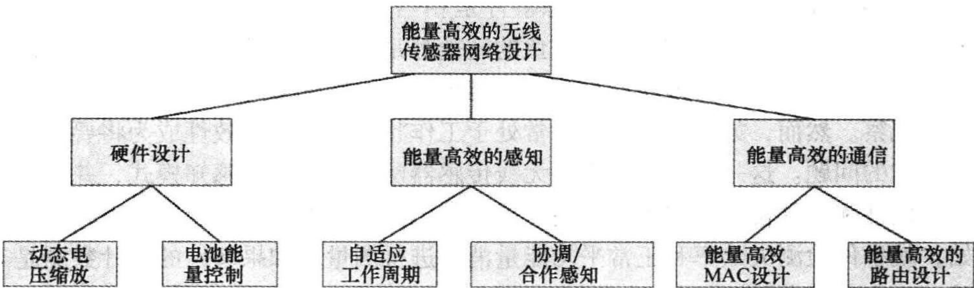


图13-1 节能模式

13.2.1 硬件能量管理

微机电系统（Microelectromechanical System，MEMS）技术、电子技术以及无线通信技术的发展促进了低成本的感知节点的发展。一个感知节点通信由一个感知单元（一个A-D转换器和一个嵌入式传感器）、一个通信单元（RF电路）、一个能量单元（电池）以及一个处理单元（带内存的处理器）组成。例如传感器节点使用来自XBOW的Mica2/Micaz节点，它由一个感知板、一个通信板、两个AA电池以及8bit微控制器组成[8]。为了定义一个无线传感器的系统结构，设计者们首先需要选择一个微控制器或者一个感知计算机[2]。

13.2.1.1 动态电压缩放比

对于一个基于CMOS的处理器，能量消耗可以划分为转换和泄漏两部分。提高微处理器能量有效性的技术之一，是利用动态电压缩放比（Dynamic Voltage Scaling，DVS）。DVS指在使用一个处理器时，根据波动来动态调整电压供应以及时间

表频率。

DVS 已经被应用于商业可用的微处理器中，它利用处理器中工作负载的变化，并且通过实现电路层的能量均衡来限制延迟。DVS 既降低了泄漏的能量消耗，也降低了转换的能量消耗。执行 DVS 的感知节点比固定电压处理器[3,4]的感知节点节约大约 60% 的能量。执行 DVS 的一个例子是 μAMPS 感知节点[5,6]。

13.2.1.2 能量资源管理

对于大部分传感器节点，电池是主要的限制因素，它抑制了传感器节点在无人监护的环境下工作数年。出现了一种趋势是在无线嵌入式感知系统[2]的能量子系统中使用超级电容器作为一个中间元件。将超级电容器与电池并联可能是解决这个问题方法的一部分。这样做转瞬即逝的能量由电容器传递，这是与电池相反的。另外一种技术，是着重于电池能量控制的感知能量资源，研究者们提出了电池友好的释放形式。

13.2.2 能量有效的无线通信

针对 WSN，已经提出了大量的能量有效的 MAC 协议和路由协议，例如为 MAC 设计的 S-MAC[13]、T-MAC[14]、STEM[12] 以及 Span[17]，和为路由设计的直接扩散 (directed diffusion)[18]、谣传路由 (rumor routing)[19] 以及 SPIN[20]。因为交替感知有一个相似的睡眠/监听机制也已经被感知 MAC 设计所采用，我们将会简要地回顾一些相关的能量有效的 MAC 设计方法。

13.2.2.1 基于竞争的 MAC

S-MAC 是无线传感器网络中一个基于竞争的节能的 MAC 协议。它利用一个简单的调度机制，允许邻居长期休眠和同步苏醒。在苏醒阶段，节点利用 CSMA/CA 通信。它显著降低了空闲监听的能量浪费，但是它也增加了延迟，因为在睡眠时期到达的数据将会排队等候下一个活动周期的到来。T-MAC[14] 试图通过引入自适应负载周期动态结束活动时间来改进 S-MAC。这也会降低空闲监听的能量消耗。STEM[12] 协议提供了一个当节点睡眠时建立通信的方式。这个协议利用两个无线电结构，数据无线电设备和唤醒无线电设备。发送节点利用数据无线电设备发送数据给目标节点。唤醒无线电节点利用极低能量的无线电设备来唤醒目标节点。如果一个节点想要建立通信，开始发送激活一个指定用户的信号。在有限的时间内，被指定用户将会被唤醒并开始通信。在参考文献 [15] 中采用了一个替代方法，允许一个传感器用一个忙标志来唤醒一个邻居而不是利用信号。然而，这也导致了一个问题：每个忙标志一定会唤醒一个节点的全部邻居，因为目标接收者的 ID 在唤醒信道中未进行编码。参考文献 [15] 的主要贡献是通过概率评估能够有选择性地唤醒先前参与通信的节点上的数据无线电设备。Span[17] 通过选择互连的一系列节点作为协调器，关闭剩下的节点，来降低多跳自组织网络的能量消耗。Span 协调器始终保持清醒，然后执行多跳包路由。

13.2.2.2 基于TDMA的MAC

TDMA-Wakeup（TDMA-W）是一个基于TDMA的调度协议[16]。在一个典型的TDMA模式中，每个节点分配一个时隙，只能在时隙内进行数据的接收和发送。在TDMA-W中，每个节点有两个时隙：发送时隙（S-slot）和唤醒时隙（W-slot）。一个节点在它的发送时隙内监听信道。如果被其他节点占用，它开始监听发送者何时占用发送时隙。当节点之间的距离超过两跳时，一个唤醒时隙能够被一个或者多个节点共享。不足的是，当在一个TDMA-W片中接收到超过一个唤醒消息时，会导致碰撞。这导致节点开始监听它的所有邻居来确定数据源。

13.2.3 能量有效的感知

尽管对于无线通信单元已经提出了许多节能模式（例如参考文献［12-14，17］），很少专门考虑感知单元。最新测量结果显示感知节点进行感知测量时要花费大量能量[11]，并且感知能量消耗与数据通信（发送和接收）消耗是相当的。再者，感知频率通常高于无线通信频率。例如，一个监测森林火灾的传感器节点仅当感知温度超过阈值时才会传送警报信号。因此，研究感知的节能模式是有必要的，换句话说，在没有活动的感知需要处理时，让感知板进入休眠状态。这一章，我们将回顾能量有效的感知技术。然后，在下面的章节，我们将关注于检测动态非合作事件的能量有效的感知。

大体上，有两种能量有效的感知方法。

1）自适应的感知占空比：在自适应感知方法中，提出了自适应地调整睡眠/感知占空比，来最好地评估/适应事件发生模式。

2）非合作的协调/合作感知（Coordinated/Collaborative sensing）：在非合作的协调/合作感知中，合作模式被提出，用于探索一个公共区域或者事件的感知的时-空坐标。

13.2.3.1 自适应的感知负载周期

WSN通常是用来监测静态或者动态事件。静态事件（例如温度和湿度）很容易测量。相反地，动态事件是非合作的，因此很难捕捉。战场上敌方飞机的移动，海洋中鲸鱼的迁徙，都是非合作事件的例子。由于它们的频繁移动很难被监测。因此，传感器只有连续地进行环境监测才能够观察到它们。然而，感知板一直处于工作状态，能量有效性会成为一个严重的问题。因此，研究感知动态非合作事件监测的节能模式是关键。

已经有了静态事件监测的几个方案，例如ELECTION[21]。在ELECTION[21]中，感知节点周期性地感知环境，睡眠时关闭传感器。它也根据测量值调整睡眠周期。为了测量时间紧急数据，仅当测量值超过了阈值时节点才会转换为无线通信模式，将数据报告给基站。

在参考文献［27］中描述了大范围环境感知网络中最优的数据收集和数据融

合算法。使用一个利用指数级后退策略的归纳模型来收集最佳数量的数据。然而，使用加倍/减半算法进行感知周期的控制是受限的。对环境改变的响应可能是缓慢的。事实上，作者也指出“在数据收集算法中的加倍和减半睡眠时间的技术，可以根据数据的动态性，调整增加和减少的力度来提高性能。”参考文献［27］的另外一个问题是所有的传感器节点同时进行数据感知。这个可以通过相邻传感器节点协作来得到提高。数据融合问题可以通过引入一个新的概念，即超感知来解决。它是基于传感器之间的自组织和协作。

在 DANCE[28] 中改进了在参考文献［27］中提及的方法，所有的测量数据和邻居节点的感知信息在决定感知调度中都会被考虑到。在 DANCE 中，如果邻居节点拥有相似的感知调度，每个传感器节点将会改变它们的感知时间。否则，如果没有邻居节点执行感知，那么传感器节点自身便会执行感知任务。因为 DANCE 将感知时间分散化，最小化了域和检测之间的延迟，所以这个方法降低了能量消耗的无效性，同时提高了感知行为的可靠性[28]。

在参考文献［25］中考虑到分布式数据流结构，因为大量传感器节点会持续地将数据传送到中央服务器。每个传感器节点的数据采样率影响到中央服务器的通信资源和计算负载。使用一个卡尔马过滤器（Kalman Filter，KF），传感器节点可以利用 KF 评估错误来自适应地调整它在规定范围内的采样率。如果当前采样率超过了这个范围，传感器节点和服务器会协商一个新的采样率。

在参考文献［26］中提出一个自适应控制机制，来改变每个传感器节点的测量频率。随着可用数据的增加感知阶段的精确性得到了提高，感知阶段的变化也会适应感知环境。然而，困难在于对现实环境的建模。

13.2.3.2 协调/合作感知

传感器网络的一个主要目标是在尽可能长的时间内保持传感器网络的覆盖范围。网络覆盖范围的问题决定了一个传感器网络的节点监测是否是良好的。过去数年，对传感器网络覆盖范围已经做了大量研究。

Meguerdichian 等人[29]定义了一个感知覆盖度量，叫做监视，可以作为一个特定传感器网络的服务质量的量度标准。一个最优的多项式时间的利用图表理论和计算几何结构的算法被提出，用来解决最佳和最优情况的覆盖范围。然而，该算法高度依赖于特定的几何结构，例如 Delaunay 三角测量和 Voronoi 图，在分布式方式中不能够进行或者有效地进行局部构建。参考文献［32］改进了参考文献［29］，提出了更加有效的分布式算法。He[52]等人提出了一个能量有效的监视系统（Energy-efficient Surveillance System，ESS），通过自适应调整系统的敏感度对能量消耗和监视行为的均衡进行了研究。

在参考文献［50］中睡眠调度问题被认为是一个限制电池寿命和感知覆盖范围的最大问题。提出了一个集中式算法和一个分布式算法，获得 k 覆盖的同时最大化了网络寿命。如果传感器密度很高，那么提出的模式能够保证指定的感知覆

盖度。

在参考文献［53］中提出了一个称为轻量级部署感知调度（Lightweight Deployment-Aware Scheduling，LDAS）的分布式调度机制。它假设感知节点没有装备GPS或者其他获取位置信息的设备。因为没有位置信息很难决定一个节点的感知区域是否完全被其他节点所覆盖，LDAS仅仅能够在静态感知中达到指定的感知覆盖范围。

在参考文献［22］中提出PEAS，它由两个简单算法组成：探测环境和自适应睡眠，它们决定了①哪个传感器节点工作以及一个苏醒的传感器节点如何决定是否回到睡眠状态；②如何动态调整感知节点的平均睡眠时间来保持一个相对稳定的苏醒率。在参考文献［51］中提出了PECAS，它是PEAS协议的扩展。它在对邻居的探寻信息的应答中公布它的剩余工作时间，因此可以预防盲点。总之，PECAS通过增加更高的信息交换能量消耗来获得比PEAS更好的覆盖范围。

在参考文献［49］中，提出了一个简单的自调度机制，称为随机独立调度（Randomized Independent Scheduling，RIS），用于延长网络寿命，同时获得一个渐近的k覆盖。RIS不需要位置信息或者距离信息，也不需要通信消耗。

在参考文献［24，30］中，提出一个节点调度模式，通过关闭传感器网络中一些过剩的节点来降低系统的全局能量消耗。它们利用概率探寻模式决定节点的关闭时间以及什么时候节点需要重新回到活动状态。为了防止邻居节点同时关闭，设计了一个基于后退的方法。

注意，在参考文献［22，24，30］中描述的模式不能够保证区域的完全覆盖。在这些模式中，如果有邻居节点处于活动状态那么该感知节点不需要进行感知。然而，睡眠节点的感知范围并不能够完全被邻居节点所覆盖。

针对密集的分簇传感器网络提出了一个能量均衡的调度（Balanced-energy Scheduling，BS）模式。BS模式的目标是将感知与通信任务的能量负载分布到簇中的所有传感器节点，延长网络寿命直到簇不能提供足够的感知覆盖范围。

在参考文献［31］中对利用睡眠和活动状态交替来节能导致的覆盖问题进行了研究。作者考虑了覆盖问题的两种类型的机制：在随机睡眠类型中，每个传感器节点保持了睡眠活动调度彼此独立，在协作睡眠类型中，传感器节点彼此协作来达到动态的睡眠调度。很明显在调度中无论采用哪一种调度策略附加了过多的控制策略，超过了一个特定的阈值，对降低周期负载所带来的好处就会开始降低。同时也表明，虽然花费了额外的控制消耗，一个协作睡眠调度是更加健壮的，而且同随机睡眠调度相比，同样数量的冗余，协作睡眠调度能够获得一个更高的占空比。

在参考文献［38］中引入了互联传感器覆盖的概念，并且开发了一个集中的近似算法，能够组建包含一个接近最优的互联传感器覆盖的拓扑。鉴于空间查询需要提取一个物理区域中感兴趣的数据，提及的贪心算法周期性地选取一条路径上的与一个已经被选择的节点互联的传感器节点，直到所给的查询区域完全被覆盖

为止。

在参考文献［37］中提出了一个有效的方法用来延长传感器网络的寿命，它将传感器节点组成一个最大数量的一系列的成功地活动的覆盖物。仅仅是当前活动的传感器节点负责监测所有的目标和传输收集的数据，其他所有传感器节点是处于一个低能耗的睡眠模式。

感知覆盖问题也扩展到了移动传感器网络。它的目标是最大化一个给定的目标区域的覆盖范围，然而它在调度时间，传感器节点最大化覆盖范围的传输距离，协议的复杂性上都是受限的。在参考文献［46］中讨论了移动传感器相对于静态传感器取得的进展，以及针对移动传感器提出的最优策略。

在参考文献［44］中提到了移动传感器网络自身调度的一个潜在的基于区域的方法。它假设每个节点上装有一个传感器，使它能够确定它的邻近节点以及障碍物的距离和方向。每个节点会因靠近它的邻近节点和障碍物而回退，因此将网络散播到整个环境中。这种方法在感知过程中既是分布式的又是可测量的，不需要环境模型、定位或者节点通信。

Heo 等人[45]提出了两个模式，叫做分布式自散布算法（Distributed Self-Spreading algorithm，DSS）和智能部署与分簇算法（Intelligent Deployment and Clustering Algorithm，IDCA）。在 DSS 中，假设传感器节点最初随机部署，它们因邻居节点施加的部分影响开始移动。在 IDCA 中，局部密度与所有节点的统一分布密度相比较。然后由节点的相对剩余能量决定它是否移动。这种算法的思想是为了在所有节点统一部署在区域中时降低节点间剩余能量的差异。

在参考文献［42］中，Voronoi 图表用于所有传感器节点初始部署以后发现覆盖漏洞。Ganeriwal 等人[43]提出了一个协议，在部署的传感器网络中利用传感器节点的灵活性修复由于能量损耗所带来的覆盖范围的损失。当节点遭到物理破坏时将会失效。

在许多提出的方法中将网络连通性问题同感知覆盖范围一起考虑，例如参考文献［33-36，39-41］。当传感器节点的传输范围至少是它的感知范围的二倍，保证 k 覆盖将会导致 k 连通[33,34]。总体上，高连通性通常能够保证高健壮性。然而，如果连通性太高，节点之间的数据冲突可能严重影响数据传输率。在参考文献［39］中提出的最近研究成果不依赖于假设传感器节点的通信范围不低于它们感知范围的二倍。Shakkottai 等人[36]考虑一个基于度的传感器网络，它由一系列有可能失效的传感器节点组成，并且研究了覆盖范围、连通性和网络直径。在参考文献［40］中提出了一个最佳的部署策略，能够达到完全覆盖和所有通信和感知范围的 2 连接。对于感知覆盖和网络连通性所提出的不同模式的更加细节的对比，读者可以参考 Wang 和 Xiao[47]。

由于 WSN 中的传感器节点完全分布式地工作，没有一个中央控制或者对网络的全局认识很难组织节点的睡眠和活动模式的转换。

为了最小化事件错过率，同时最小化和平衡节点的能量消耗，需要一个分布式模式让传感器节点进入睡眠状态，并且在需要时唤醒传感器节点。这一章的余下部分主要关注的是各种分布式交替感知模式，包括对固定和自适应模式的技术性能分析以及它们对于动态非合作事件网络覆盖的意义。特别是，在传感器网络中能量有效性和动态非合作事件覆盖范围的均衡被描述为一个连接最优化问题。首先我们由单一传感器节点的固定的睡眠监听时间模式开始，测量其事件错过率和正常的平均能耗的执行结果。如果事件的统计数据是未知的，提出了一个单一传感器节点的自适应模式来适应有感知/测量数据的睡眠间隔。固定模式和自适应模式都扩展到了多个节点，除了能量有效性也将整个网络区域的覆盖范围也考虑进去。

13.3 交替感知模式

尽管先前的工作尽量最小化数据收集层的能量消耗，它们都没有强调非合作事件的感知机制。动态非合作事件在时间上可能是持续的或者分离的。检测动作非合作事件需要考虑到使用一个决策反馈来决定传感器节点的感知行为，例如占空比，而不是记录持续的数据流或者数据集合。



图 13-2 另外一个感知模式

本文中，传感器网络中采用交替感知模式来节省能量。这种模式的一个例子如图 13-2 所示。在监听时间内，感知节点打开它们的感知板进行测量。在睡眠时间内，传感器节点转换到关闭状态来节省能量。尽管这些模式被广泛采用，如何设计监听和睡眠时间间隔的分配来监测动态非合作事件仍然是一个悬而未决的问题。

本文中采用下面的标记：

L——监听时间，单位为 s；

S——睡眠时间，单位为 s；

λ——非合作事件的泊松率；

E_S——感知一次事件的单位能量消耗；

P_S——睡眠时期的平均能量消耗；

P_L——监听时期的平均能量消耗。

交替感知模式的设计可以被描述成是一个连接最优问题：

$$\mathrm{argMin}_{L,S}\ J \tag{13-1}$$

式中，

$$J = \{E[R_{miss}]^{+}E[\overline{P}]\} \tag{13-2}$$

期望的事件错过率，$E[R_{miss}]$，由下式给出

$$E[R_{miss}] = E\left\{\frac{\lambda S}{\lambda S + \lambda L}\right\} = \frac{E[S]}{E[S] + E[L]} \tag{13-3}$$

正常化的平均能耗，$E[\overline{P}]$由下式给定：

$$E[\overline{P}] = \frac{E[L] + \frac{P_S}{P_L}E[S]}{E[L] + E[S]} \tag{13-4}$$

假设测量数据对事件的监测是精确的。换句话说，在传感器节点感知范围内并且节点处于工作状态，无论何时有事件发生都会被准确无误地检测出来。执行索引J，描述了事件错过率（由于传感器睡眠）和能量消耗之间基本的均衡。连续的监听和睡眠时间间隔需要做到最小化预计的事件错过率和正常化的平均能量消耗的总合。如果网络的最终状态是可知的，可以使用一个动态设计方法来解决这个最优化问题。然而，实际上并非如此，其他次优的设计方法（启发式方法）也是需要的。

这里考虑两种类型的交换感知模式：固定时间（Fixed Timer，FT）模式和自适应时间（Adaptive Timer，AT）模式。在一个 FT 模式中，睡眠时间和监听时间是提前设置好的，一旦设置好了便是固定的。在一个 AT 模式中，每个节点试图根据（估计的）事件发生频率和一些用户指定的设计参数来动态调整它们的睡眠调度。

当事件的统计数据是可知的或者能够被提前预测时，FT 模式是可用的。AT 模式比 FT 模式更加灵活，并且拥有评估事件统计信息和调整事件错过率与能耗消耗之间均衡的能力，这在实际中是可取的。

本章，我们提出了下面的 AT 模式，模式中的睡眠时间根据 AIMD 规则[55]改变

$$S(k) = \begin{cases} S(k-1) + \delta S(k) & I(L(k-1)) = 0 \\ \beta(k) S(k-1) & I(L(k-1)) = 1 \end{cases}$$

这里 $\delta S(k) > 0$ 是增加步长，$0 < \beta(k) < 1$ 是在一个时间步长 k 的降低因素。它们都是设计参数。I（.）是一个指示函数，定义为

$$I(t) = \begin{cases} 1 & \text{如果在 } t \text{ 时间内检测到事件} \\ 0 & \text{否则} \end{cases}$$

所提出的 AT 模式的本质是动态探测事件的统计数据，据此调整睡眠间隔。这两个设计参数提供了事件错过率与能量消耗之间的保持均衡的能力。例如，较大的 β 能够节省更多能量，但是也增加了错过率。

读者可能注意到所提出的 AT 模式利用和 TCP 拥塞控制中使用了同样的原

则[56]。确实，TCP拥塞控制的本质是测量路径吞吐量（拥塞层）并相应地调整数据传输。同样的，所提出的AT模式的本质是估计事件发生的趋势，并且相应地适应睡眠调度。

我们注意到所提出的AIMD模式将会使网络处于最佳的运行状态点，即有效性与公平性获得最佳的折中[56]。有效率是平均事件错过率的相反数，然而公平性由传感器节点的平均能量消耗来反映。值得一提的是，正如在参考文献［56］中所证明的，不像AIMD，许多先前的方法采用一个潜在的后退策略，倍加/倍减(MIMD)，可能并不能达到最佳的运行状态点。

尽管所提出的AIDM模式很简单，它在现实的传感器网络中表现良好。当前的传感器都是用非常有限的内存空间存储程序。例如，MICA2节点仅有8KB的内存[8]。再者，传感器通常是具有有限的计算能力，并且很难进行故障排除。因此，在无线传感器网络中更可能采用简单的机制。

13.4 性能分析

本节采用不连续事件模拟来比较FT模式与所提出的AT模式。OPNET[59]被选作模拟工具。纯理论的FT模式的性能可以在参考文献［55］中找到，在此处省略。总体上，采用FT模式的事件错过率与能量消耗依赖睡眠监听时间间隔。因此，最小化能量消耗和错过率，网络设计者能够利用错过率和能量均衡曲线为传感器网络选择一个适当的操作点来最好地符合应用需求。

在这次模拟学习中，非合作事件根据具有一个泊松率 λ 的泊松过程产生，λ 随时间改变。在头20s中 $\lambda=1$，然后从20~40s增加到10，然后从40s到100s降低到0.1，然后从那以后返回到1。改变事件率的目的是建立一个真实的（动态的）环境，然后将FT模式与提出的AT模式进行比较。

事件错过率和正常的平均能耗消耗利用两个FT模式（相应的 $S=1$，$L=1$ 和 $S=0.1$，$L=1$）和提出的AT模式，它们如图13-3所示。我们发现不恰当设置的FT模式（$S=1$，$L=1$）的表现是不令人满意。一个具有更好设置的FT模式（$S=0.1$，$L=1$）能够获得更好的表现。然而，在FT模式中设置合适的参数需要首先知道事件过程。相反地，AT模式不需要事先知道事件过程。更加重要的是，所提出的AT模式在事件错过率和正常平均能量消耗两个方面都取得比FT模式更好的表现。

在图13-4中显示了不同参数设置的AT模式的事件错过率和正常平均能量消耗。很显然所提出的AT模式在大范围的参数设置上表现良好。它也证明了所提出的AT模式的灵活性。

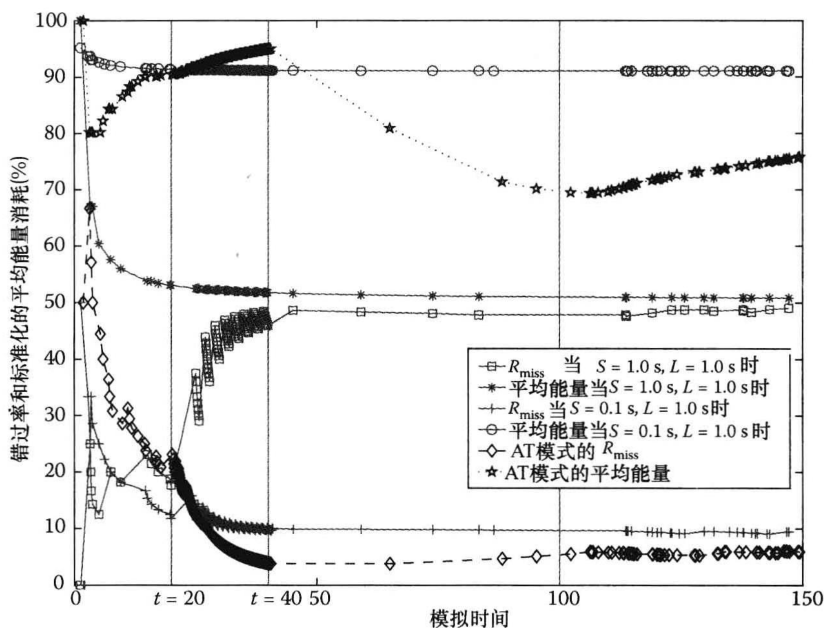


图 13-3　使用两个 FT 模式和提出 AT 模式（其中 $\delta S=0.1$，并且 $\beta=0.5$）事件错过率和标准化的平均能量消耗

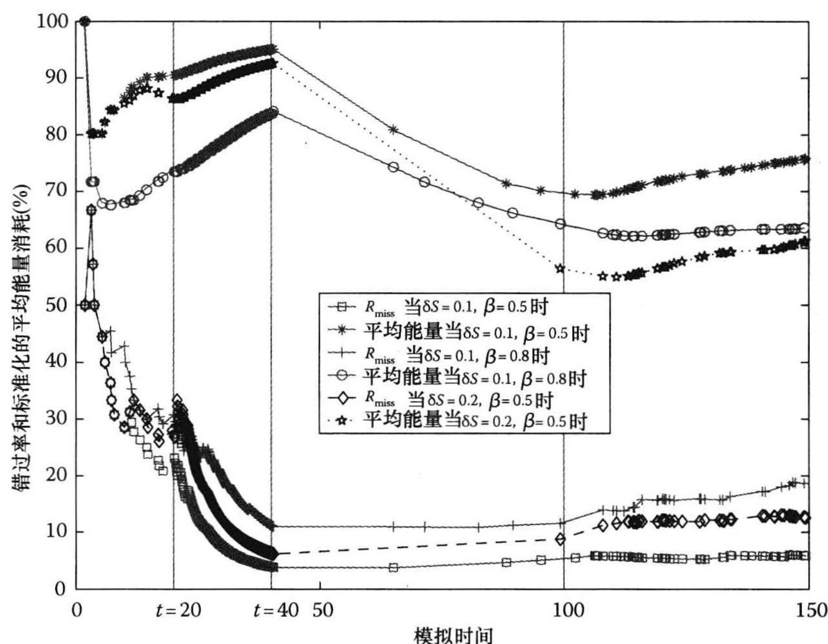


图 13-4　使用不同参数设置 AT 模式的事件错过率和标准化的平均能量消耗

13.5　网络充分覆盖范围

在先前的章节中，研究了单一节点监测动态非合作事件的FT模式和AT模式。然而，实际应用中，经常需要监测整个无线传感器网络覆盖区域。例如一个WSN可能部署来监测敌方活动的战场。因此，需要研究各种网络覆盖模式下FT模式和AT模式的表现。

总体上，没有中央控制器很难协调感知节点。分布式控制需要引入大的代价损失。再者，对于不可以或者不方便协调传感器节点睡眠的应用，随机睡眠是仅有的最优机制。因此，传感器节点工作中将会采用一个异步模式，传感器节点不需要彼此同步，每个传感器节点根据在13.3节中所提出的AT模式独立地决定它们的睡眠调度。与先前在网络覆盖上所做的工作相反，我们将自适应占空比设计与地理覆盖结合起来，采用最小假设和信号需求相结合的方式来得到结果。事件错过率由事件发生和每个单一传感器的自适应占空比两方面来决定，而不是单纯地由传感器节点的地理位置决定。

13.5.1　理论结果

我们的分析中采用下面的标记：

(X, Y)：传感器节点的二维坐标（位置）；

N：网络中感知节点数；

r：一个感知节点的感知半径；

A：网络的区域。

假设随机的非合作事件根据一个泊松率为λ的泊松分布发生。事件的位置随机选择，并且假设统一分布于整个网络中。

定理13.1

假设有N个传感器节点统一分布在一个WSN区域A中，每个传感器节点有感知半径r，错过率（错过一个随机事件的概率）如下所示。

$$R_{miss} = e^{-\frac{E[L]}{E[S]+E[L]}\frac{N}{A}\pi r^2} \tag{13-5}$$

证明　一个随机事件发生在N_0个节点的感知范围内的概率是

$$P_0 = \frac{(\pi r^2 N/A)^{N_0}}{N_0} e^{-\pi r^2 N/A} \tag{13-6}$$

预期的一个随机事件范围内的传感器节点数量是

$$E[N_0] = \frac{N}{A}\pi r^2 \tag{13-7}$$

预期的事件发生在其感知范围内，并且自身处于监听状态的节点数量是

$$E[N_0^L] = \frac{E[L]}{E[S]+E[L]}\frac{N}{A}\pi r^2 \tag{13-8}$$

那么，一个随机事件被某个传感器节点感知到的概率是

$$1-e^{-\frac{E[L]}{E[S]+E[L]}\frac{N}{A}\pi r^2}$$

这个定理说明，事件错过率是每个单元区域内平均节点数量和睡眠调度的函数。执行分离事件模拟来评估在不同单元区域平均节点数量下 FT 模式与 AT 模式的表现。

需要指出的是这里仅得到了 1 覆盖结果[55]。在最近的一项独立工作中[48]，得到了异步随机感知中的静止的 k 覆盖率和预期的覆盖时间。对于监视应用，在参考文献［48］中也给出了检测概率和检测延迟分布。

13.5.2　模拟结果

WSN 中有 N 个传感器节点统一分布于网络覆盖区域中。每个节点有感知范围为 r 的一个二进制事件探测器。假设网络覆盖了一个 100m × 100m 的区域。一个采样传感器位置地图（$N=100$）如图 13-5 所示。注意，这里由于实际的限制，节点的覆盖并没有达到最优。例如，感知节点由无人飞机部署。

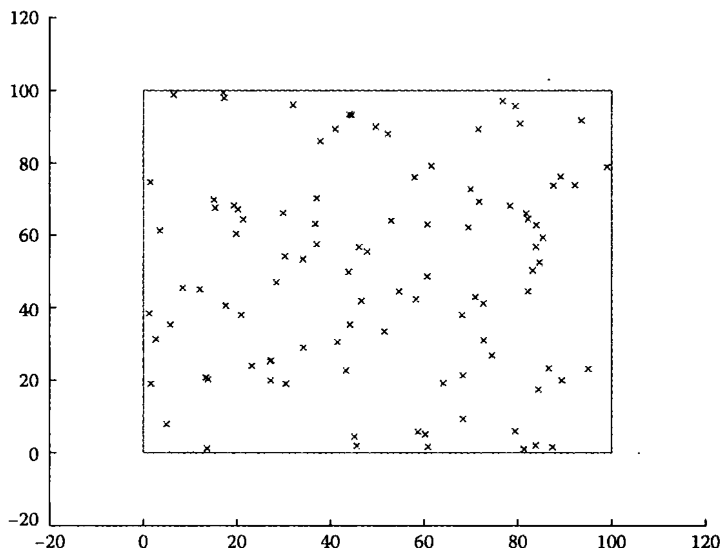


图 13-5　样本 WSN：节点位置

在 13.4 节中假设随机的非合作事件依照一个泊松分布发生。事件发生的位置是随机选择的，假设统一分布于网络中。换句话说，一个随机事件的坐标，X 和 Y，都是统一分布在［0，100］之间。

为了评估 FT 模式的网络覆盖范围，假设所有节点有相同的能量预算，事件错过率作为衡量标准。一个错过的事件意味着一个事件发生了，但是在其范围内没有感知节点或者感知节点处于睡眠状态。

FT模式的参数是 $S=0.1$ 和 $L=1$，相应地，所提出的AT模式的参数是 $\delta S=0.1$ 和 $\beta=0.5$。在不同的节点数量和不同的感知半径下，将FT模式与提出的AT模式进行比较。在图13-6（不同的 N）和图13-7（不同的 r）中对结果进行了总结。很显然，在所有实验上AT模式都比FT模式有更小的事件错过率。图13-6也显示了，当 N 增大时，由于节点密度增加FT模式和AT模式之间的差别逐渐减小。当感知范围 r 很小时，地理覆盖范围决定了事件错过率。因此，FT模式和AT模式之间有很小的差别。随着 r 的增大，感知模式在事件覆盖中开始扮演了一个重要的角色，FT模式和AT模式之间的差别开始增大。

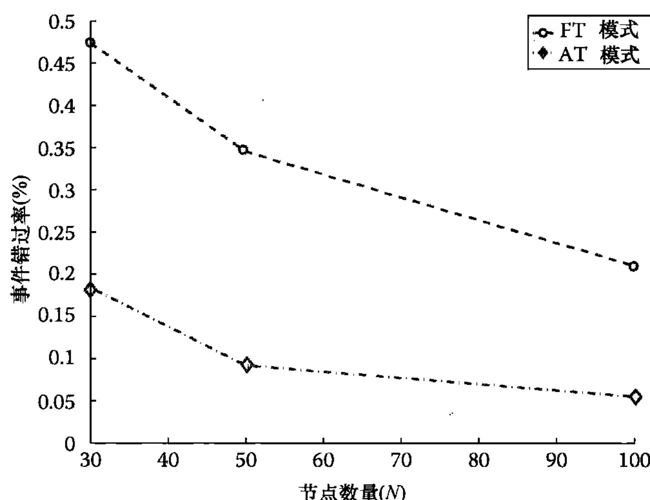


图13-6　网络中的事件错过率与传感器节点的数量比较（r 是固定的，并且 $r=15\text{m}$）

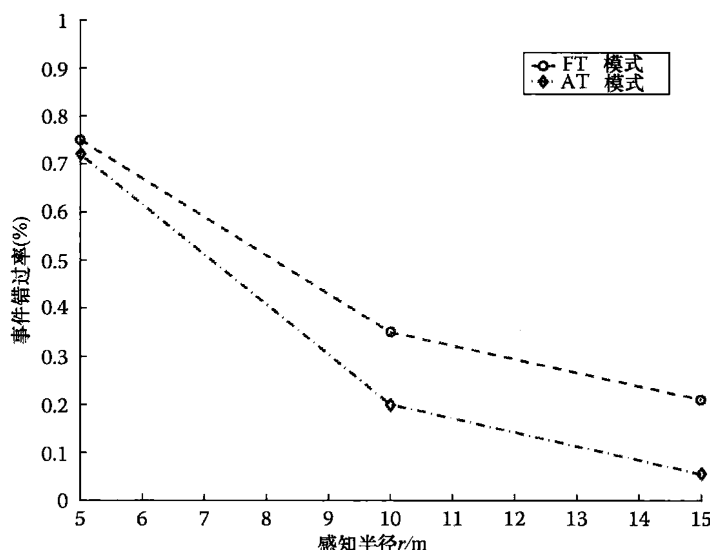


图13-7　事件错过率与感知半径比较（N 是固定的，并且 $N=100$）

13.6 尚未解决的问题和争议

在先前的大部分无线传感器网络能量有效性的设计方法中，并没有把感知和通信能量明确地区分开来。通常假设所有感知节点处于睡眠状态或活跃状态。甚至当进行感知通信行为，能量消耗分开考虑时，它们的睡眠调度也不是以整体的方式来对待。再者，许多情况下，当设计睡眠调度时不同感知应用的影响没有考虑。

先前的针对网络覆盖范围的工作主要关注的是几何分析。尽管先前已经多次尝试了将感知半径和通信半径都考虑进去，但并未分析感知调度和传输负载的动态性。

无线传感器网络的最终设计目标有两个在某种程度上互相冲突的方面：一方面，需要最小化网络覆盖范围内的事件错过率；另一方面，感知节点具有有限的资源，例如能量，它们需要达到一个最长的生命周期。为了满足这些需求，感知、无线通信和数据处理算法（例如最近提出的密集追踪压缩感知[57,58]）需要进行结合设计来获得最好的折中。再者，需要将能量资源模型考虑进去来精确地表示一个传感器节点的能量流。最后，当整合设计中的感知调度的动态性与通信需求时，重要的是从网络覆盖点的角度来考虑，而不是从每个单独的传感器节点来考虑。在图 13-8中强调了上面的建议。

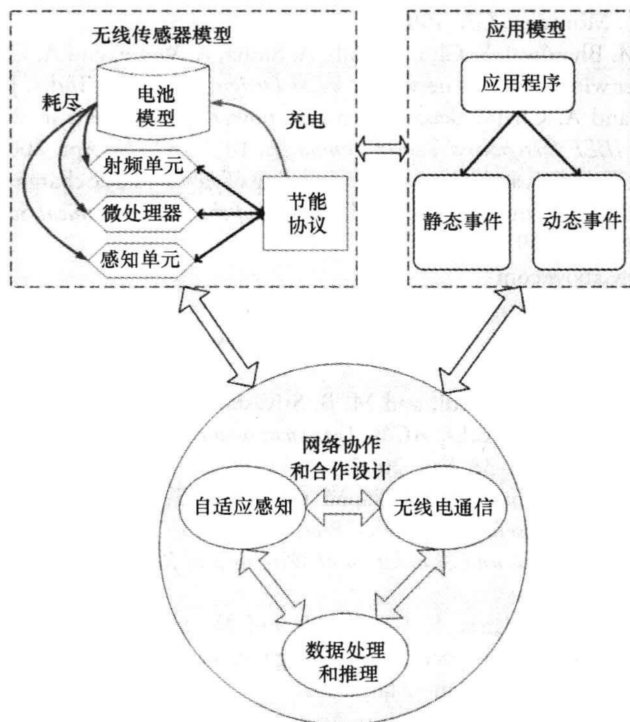


图 13-8　完整的 WSN 设计

13.7 总结和对未来工作的展望

本章中，考虑到能量有效性与故障容忍事件监测之间的折中采用灵活的动态感知调度，研究了动态改变 FT 和 AT 模式中的参数对能量有效感知的影响。具体的模拟研究提供了对实际部署的指导，同时也描述了网络覆盖结果。

注意，当在实时应用中，如果能够系统地制定增加步长和降低因素，所提出的模式才可能得到推广。再者，将我们的提议与其他方法进行比较也是有意义的，例如在参考文献［54］中，跟踪指定的目标，这也将成为未来需要努力的方向之一。

参考文献

1. I. F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, A survey on sensor networks, *IEEE Communications Magazine*, 40(8), 102–114, Aug. 2002.
2. P. H. Chou and C. Park, Energy efficient platform designs for real-world wireless sensing applications, *IEEE International Conference on Computer Aided Design*, pp. 912–919, San Jose, CA, 2005.
3. A. Hac, *Wireless Sensor Network Designs*, Wiley, New York, 2003.
4. T. Pering, T. Burd, and R. Broderson, The simulation and evaluation of dynamic voltage scaling algorithms, *Proc. Int. Symp. on Low Power Electronics and Design*, pp. 76–81, Monterey, CA, 1998.
5. R. Min, M. Bhardwaj, S. Cho, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan, Low power wireless sensor networks, *VLSI Design*, Bangalore, India, Jan. 2001.
6. A. Sinha and A. Chandrakasan, Dynamic power management in wireless sensor networks, *IEEE Design and Test of Computers*, 18(2), 62–74, Apr. 2001.
7. M. Doyle, T. Fuller, and J. Newman, Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell, *Journal of the Electrochemical Society*, 140(6), 1526–1533, June 1993.
8. http://www.xbow.com
9. S. Roundy, P. K. Wright, and J. M. Rabaey, *Energy Scavenging for Wireless Sensor Networks: With Special Focus on Vibrations*, Kluwer Academic Publishers, Boston, MA, 2003.
10. A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, Power management in energy harvesting sensor networks, *ACM Transactions on Embedded Computing Systems*, 6(4), article 32, pp. 1–38, Sep. 2007.
11. G. Anastasi, A. Falchi, A. Passarella, M. Conti, and E. Gregori, Performance measurements of motes sensor networks, *Proc. of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Venice, Italy, Oct. 2004.
12. C. Schurgers, V. Tsiatsis, S. Ganeriwal, and M. Srivastava, Optimizing sensor networks in the energy-space-density design space, *IEEE Transactions on Mobile Computing*, 1(1), 70–80, Jan.–Mar. 2002.
13. W. Ye, J. Heidemann, and D. Estrin, An energy-efficient MACprotocol for wireless sensor networks, *Proc. IEEE INFOCOM*, New York, pp. 1567–1576, June 2002.

14. T. V. Dam and K. Langendeon, An adaptive energy-efficient MAC protocol for wireless sensor networks, *ACM SenSys'03*, pp. 171–180, Los Angeles, CA, Nov. 2003.
15. M. Miller and N. Vaidya, A MAC protocol to reduce sensor network energy consumption using a wakeup radio, *IEEE Transactions on Mobile Computing*, 4(3), 228–242, May–Jun. 2005.
16. M. Zohaib and T. M. Jadoon, Comparison of S-MAC and TDMA-W protocol for energy efficient wireless sensor networks, *International Conference on Emerging Technologies*, pp. 486–492, Peshawar, Pakistan, Nov. 2006.
17. B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks, *ACM Wireless Networks Journal*, 8, 481–494, Sep. 2002.
18. C. Intanagonwiwat, R. Govindan, and D. Estrin, Directed diffusion: A scalable and robust communication paradigm for sensor networks, *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networks (MobiCom'00)*, Boston, MA, Aug. 2000.
19. D. Braginsky and D. Estrin, Rumor routing algorithm for sensor networks, *WSNAŠ02*, Atlanta, GA, Sep. 2002.
20. J. Kulik, W. Heinzelman, and H. Balakrishnan, Negotiation-based protocols for disseminating information in wireless sensor networks, *Wireless Networks*, 8(2/3), 169–185, 2002.
21. S. Begum, S. Wang, B. Krishnamachari, and A. Helmy, ELECTION: Energy-efficient and low-latency scheduling technique for wireless sensor networks, *The 29th Annual IEEE Conference on Local Computer Networks (LCN)*, Tampa, FL, Nov. 2004.
22. F. Ye, G. Zhong, S. Lu, and L. Zhang, PEAS: A robust energy conserving protocol for long-lived sensor networks, *Proc. IEEE Int. Conf. Network Protocols (ICNP)*, Paris, France, 2002.
23. J. Deng, Y. Han, W. Heinzelman, and P. Varshney, Balanced-energy sleep scheduling scheme for high-density cluster-based sensor networks, *Computer Communications*, 28(14), 1631–1642, Sep. 2005.
24. D. Tian and N. D. Georganas, A node scheduling scheme for energy conservation in large wireless sensor networks, *Wireless Communications and Mobile Computing Journal*, 3(2), 271–290, Mar. 2003.
25. A. Jain and E. Y. Chang, Adaptive sampling for sensor networks, *Proc. International Workshop on Data Management for Sensor Networks*, 72, 10–16, Toronto, Canada, 2004.
26. A. D. Marbini and L. E. Sacks, Adaptive sampling mechanisms in sensor networks, *London Communications Symposium*, London, U.K., 2003.
27. R. Dantu, K. Abbas, M. ONeill II, and A. Mikler, Data centric modeling of environmental sensor networks, *Proc. IEEE Globecom*, pp. 447–452, Dallas, TX, 2004.
28. J. Lee, D. Lee, J. Kim, W. Cho, and J. Pajak, A dynamic sensing cycle decision scheme for energy efficiency and data reliability in wireless sensor networks, *Lecture Notes in Computer Science*, no. 4681, pp. 218–229, 2007.

29. S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, Coverage problems in wireless ad-hoc sensor, *IEEE INFOCOM*, pp. 1380–1387, Anchorage, AK, 2001.
30. D. Tian and N. D. Geoganas, A coverage-preserving node scheduling scheme for large wireless sensor networks, *First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA)*, Atlanta, GA, 2002.
31. C. Hsin and M. Liu, Network coverage using low duty-cycled sensors: Random & coordinated sleep algorithms, *3rd International Symposium on Information Processing in Sensor Networks (IPSN)*, Berkeley, CA, 2004.
32. X. Li, P. Wan, and O. Frieder, Coverage in wireless ad hoc sensor network, *IEEE Transactions on Computers*, 52(6), 753–763, June 2003.
33. X. Wang, G. Xing, Y. Zhang, C. Lu, R. Pless, and C. Gill, Integrated coverage and connectivity configuration in wireless sensor networks, *ACM SenSys'03*, Los Angeles, CA, Nov. 2003.
34. H. Zhang and J. C. Hou, Maintaining sensing coverage and connectivity in large sensor networks, *NSF International Workshop on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless and Peer-to-Peer Networks*, Fort Lauderdale, FL, Feb. 2004.
35. J. Lu, L. Bao, and T. Suda, Probabilistic self-scheduling for coverage configuration in wireless ad-hoc sensor networks, *International Conference on Sensing Technology (ICST)*, Palmerston North, New Zealand, 2005.
36. S. Shakkottai, R. Srikant, and N. Shroff, Unreliable sensor grids: Coverage, connectivity and diameter, *Proc. of IEEE INFOCOM*, San Francisco, CA, 2003.
37. M. Cardei, M. Thai, Y. Li, and W. Wu, Energy-efficient target coverage in wireless sensor networks, *Proc. of IEEE INFOCOM*, Miami, FL, 2005.
38. H. Gupta, S. Das, and Q. Gu, Connected sensor cover: Self-organization of sensor networks for efficient query execution, *MobiHoc'03*, Annapolis, MD, 2003.
39. C. Huang, Y. Tseng, and H. Wu, Distributed protocols for ensuring both coverage and connectivity of a wireless sensor network, *ACM Transactions on Sensor Networks*, 3(1), article 5, pp. 1–22, Mar. 2007.
40. X. Bai, S. Kumar, D. Xuan, Z. Yun, and T. H. Lai, Deploying wireless sensors to achieve both coverage and connectivity, *MobiHoc'06*, Florence, Italy, 2006.
41. Z. Jiang, R. Kline, J. Wu, and F. Dai, A practical method to form energy efficient connected *K*-coverage in wireless sensor networks, *ICDCSW'06*, Lisboa, Portugal, 2006.
42. G. Wang, G. Cao, and T. La Porta, Movement-assisted sensor deployment, *IEEE INFOCOM*, Hong Kong, China, 2004.
43. S. Ganeriwal, A. Kansal, and M. B. Srivastava, Self aware actuation for fault repair in sensor networks, *IEEE International Conference on Robotics and Automation*, New Orleans, LA, 2004.
44. A. Howard, M. J Mataric, and G. Sukhatme, Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem, *6th International Symposium on Distributed Autonomous Robotics Systems*, Fukuoka, Japan, 2002.
45. N. Heo and P. K. Varshney, An intelligent deployment and clustering algorithm for a distributed mobile sensor network, *IEEE International Conference on Systems,*

Man and Cybernetics, Washington, DC, 2003.

46. N. Bisnik, A. A. Abouzeid, and V. Isler, Stochastic event capture in mobile sensor networks subject to a quality metric, *MobiCom'06*, pp. 89–109, Los Angeles, CA, Sep. 2006.
47. L. Wang and Y. Xiao, A survey of energy-efficient scheduling mechanisms in sensor networks, *Mobile Networks and Applications*, 11, 723–740, 2006.
48. C. Hua and T. P. Yum, Asynchronous random sleeping for sensor networks, *ACM Transactions on Sensor Networks*, 3(3), article 15, pp. 1–25, Aug. 2007.
49. S. Kumar, T. Lai, and J. Balogh, On K-coverage in a mostly sleeping sensor network, *Mobicom 04)*, Philadelphia, PA, 2004.
50. P. Berman et al.,Efficient energy management in sensor networks, in *Ad Hoc and Sensor Networks*, Nova Science Publishers, New York, 2005.
51. C. Gui and P. Mohapatra, Power conservation and quality of surveillance in target tracking sensor networks, *Mobicom 04*, Philadelphia, PA, 2004.
52. T. He et al., Energy-efficient surveillance system using wireless sensor networks, *MobiSys 04*, Boston, MA, 2004.
53. K. Wu et al., Lightweight deployment-aware scheduling for wireless sensor networks, *ACM/Kluwer Mobile Networks and Applications (MONET)*, 10(6), 837–852, 2005.
54. S. Pattem, S. Poduri, and B. Krishnamachari, Energy-quality tradeoffs for target tracking in wireless sensor networks, *Lecture Notes in Computer Science*, no. 2634, pp. 32–46, Springer-Verlag, 2003.
55. L. Qian, A. Quamruzzman, and J. Attia, Energy efficient sensing of non-cooperative events in wireless sensor networks, *The 40th Annual Conference on Information Sciences and Systems (CISS)*, pp. 93–98, Princeton, NJ, Mar. 2006.
56. D. Chiu and R. Jain, Analysis of the increase and decrease algorithms for congestion avoidance in computer networks, *Computer Networks and ISDN Systems*, 17, 1–14, 1989.
57. E. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. on Information Theory*, 52(2), 489–509, Feb. 2006.
58. R. Baraniuk, Compressive sensing, *IEEE Signal Processing Magazine*, July 2007.
59. http://www.opnet.com

第 14 章　无线传感器网络的移动性

无线传感器网络中采用移动的感知节点和 sink 节点相对于静态设置有几个优势。移动的感知节点可以用来减少覆盖漏洞，能够跟踪被监测事件，或者当需要时聚集到监测区域中感兴趣的区域来支持更加精确地测量。另一方面，移动 sink 节点能够移动到那些报告事件的区域（这样可以降低通信距离），当多跳传输不可能的时候，能够移动到一个远处节点的附近，或者能够平衡中继感知节点上的负载，以此来均衡网络中的能量消耗。本章中，我们将详细地讨论对于感知节点和 sink 节点移动性的不同解决方法，它们的优势与不足，概述它们对于媒体访问控制（Medium Access Control，MAC）或者路由等问题的影响。

14.1　概述

几年前当无线传感器网络开始引起人们的关注时，大部分应用和主要的研究关注的是完全静态的场景。然而，人们很快认识到研究感知节点和 sink 节点的移动性可能会带来几个优势。通常假设感知节点有限的能量供应不应该浪费在节点移动上。然而，许多应用场景中，感知节点到处移动并未消耗自身的能量，而是由不同的“移动代理”运载；它们被安装到野生动物身上来监测它们的栖息地，安装到车上来增加移动安全性，或者漂浮在水面上来监测石油泄漏。当感知节点的移动不受控制时，随着它们的感知网络拓扑动态改变，保证感知数据的可用性、定位和有效的聚集是一个挑战。另一方面，受控的感知移动，也就是说，允许特定的感知节点根据一个特定算法自主移动，相对于采用一个静态的网络结构的场景也会有很多优势。因此，移动的感知节点可以用于减少覆盖漏洞，跟踪被监测到的事件，聚集于人们最感兴趣的区域，或者能够作为一个稀疏连接环境中的中继节点。

与感知节点相反，sink 节点可能有相当大的能量供应，它可能装有可再充电的或者可更换的电池，利用它们的能量来支持移动性，因此通常是更加可靠的。应用移动 sink 节点可以达到以下几个目标。一方面，不时地移动 sink 节点能够改变数据聚集路径，对网络的负载均衡是有益的。在单跳网络中，能够大大减少远离 sink 节点的感知节点数量，同时在多跳网络中，对于 sink 节点的邻居也是一样的。在这两种情形下，sink 节点的迁移有助于减轻网络中的能量消耗。sink 节点的迁移是随机的，但是也有算法监测感知节点的能量水平，使得 sink 节点避开几乎是空的区域。

另一方面，sink 节点移动靠近感知节点，降低通信距离，能够提高能量有效

性。在非时间紧急事件中，感知节点可能缓存它们读取的数据一段时间，然后仅当 sink 节点经过它时将数据交付给 sink 节点。这可能发生在一个随机的时刻，当 sink 节点随机移动或者每隔一定的时间，以及当 sink 节点服从一个预定义的聚集路径时。利用移动 sink 节点也会有这样的好处，网络不再需要完全连接；移动 sink 节点能够从稀疏分布的区域中收集感知数据，在稀疏分布区域中多跳路由是无法保证的。然而，有些应用的数据传递是时间紧急的；当在监测区域中的某块区域检测到一个异常事件时，需要立即向 sink 节点报告。在这种情况下，移动的 sink 节点也是有用的，因为它们能够自适应地靠近当前事件，因此能够减少通信路径的长度。

当采用移动的感知节点或者 sink 节点时，需要考虑几个重要因素。移动节点如何以一种能量有效的方式通知其他节点它的新位置？我们如何根据新位置重新设置通信路径？节点移动性对其他能量优化技术有什么影响，例如如何分簇或者数据聚合？所有上面提到的感知节点和 sink 节点移动性的方面将会在本章中利用特定的应用场景和用户实例来进行详细的探讨。

本章的余下部分的内容如下。首先，我们介绍了受控的和不受控的感知节点移动性的解决方法。然后，我们研究了为什么采用移动 sink 节点是非常有用的，以及对随机的、可预测的、受控的、自适应的 sink 节点移动性的解决策略。紧接着，我们研究了虚拟移动性，网络节点没有物理移动，而是节点之间不同功能的互换，这样所形成的场景和传统的移动场景是相同的。本章最后，考虑到媒体访问控制（MAC）或者路由，我们分析了感知节点和 sink 节点移动性的不同后果。

14.2 传感器移动性

传感器的移动性被广泛认为是 WSN 部署和管理复杂性的另外一个方面。同时，传感器的移动性也会引入另外一个程度的控制，使得一个传感器网络更加灵活和可重配置，达到兴趣最优均衡。

根据传感器节点如何移动和为什么移动，有两种主要的传感器节点移动（见表 14-1）。如果传感器节点仅仅是因为外部环境移动（例如水中或者空气中的漂浮传感器节点），问题不是如何移动它们，而是如何控制这种内在移动性，以此来管理和操作网络。在这种情况下，移动性是一个需要仔细处理的挑战。另一方面，如果有可能并且有能力移动传感器节点，或者至少有一些（直接或者间接的）对于传感器节点移动的控制，传感器节点移动性可以作为开发和应用的一个新工具，在某种程度上能够管理和改进网络。移动传感器节点能够依靠它们自身移动（例如通过轮子、机器腿和迷你火箭），或者绑定到一些运输机上（例如机器人和交通工具）。有移动能力的传感器能够以一个自我调整的方式移动，或者能够被网络管理者或者应用本身控制。在受控移动和非受控移动之间有一个过渡区域，对于绑定到一些“第三组织”传输机（例如动物和非受控交通工具）的传感器节点，它们的

移动方式依赖于传输机。这种情况下，网络管理者或者应用对于传感器节点的移动性可能只有很少的影响。

表 14-1 传感器移动性类型和任务

传感器移动性	摘 述	处理内在的移动
不受控的	传感器根据一些外界环境因素（例如，漂浮在水面上）被动移动	处理内在的移动[20]
受控的（分布式的或集中式的）	传感器自身主动移动（例如，通过轮子，机器腿等）。它们以一种自调整的方式（分布式的）移动或者能够被网络管理员或者应用控制（集中式的）。	提高或者增强网络覆盖率；保证连通性；响应环境的改变（例如，目标检测和跟踪）[14]；拓扑控制；提高数据读数或者在关键区域采取行动的质量[8,26,27]；
第三组织	传感器被安装到一些传输设备（如车辆或个人）四处移动。	利用传输装置[32]

当我们将受控移动作为改进网络操作的一个方式时，可能有不同的目标。移动传感器迟早能够改变或者保证网络覆盖范围或者确保整个感知区域的连通性。另一方面，重定位感知节点也可以用于自适应响应环境的改变或者聚集于人们感兴趣的区域，可能也会随着时空改变而改变。当跟踪意味着是对于被监测事件或者目标的物理跟踪时，目标监测和跟踪的目标可以作为后者的一个实例。

如果一些移动传感器能够作为中继节点，帮助数据传递，那么感知区域的连通性能够得到有效的提高。主要的思想是引导这些移动中继节点填充网络中的通信“漏洞”，也就是说，移动到两个距离比它们的最大通信半径还要远很多的结点之间。这样的移动服务会引入额外的跳数来降低其他节点的传输半径。

当我们提到拓扑控制时，我们的意思是移动传感器应该对自己定位，并且根据选择的感知部署计划，根据当前的静止传感器节点拓扑和覆盖需求调整它们的传输能量[35]。当覆盖范围非常重要时，大部分工作集中于重定位传感器的算法来获得扩大的覆盖区域的静态结构。在参考文献［15］中，对依赖于传感器移动的覆盖范围的动态方面进行了详细的研究。

再者，如果我们考虑的是非同构的传感器网络，也就是说，感知区域有带有不同功能设备的不同类型的传感器节点，移动传感器节点的任务的角色是不同的。我们很容易想到，那些可移动的传感器节点会更加昂贵，体积更大，需要更多的能量和维护代价。因此，一个可行的方案是仅有一部分节点是更加先进的移动传感器节点，用于执行特殊的任务。例如，考虑在一个无线传感器网络应用中，部署大量静态节点来监测一大片区域，收集来自整个区域的基本信息。如果传感器节点的数据显示在一个指定时间和指定地点有用户感兴趣的事件发生，通知精密的移动节点去访问那个区域以及采取进一步行动（例如实施更加先进或者精确的测量和照相）。

14.2.1 非受控移动性

当我们论及非受控传感器节点的移动时，我们经常假设传感器节点不是静态

的，而是因为一些外部环境效力而被动移动。例如，在空气中或者海洋中的传感器节点，因为空气或者海洋的作用力移动，渐渐远离它们最初被部署的位置。在参考文献 [20] 中，我们能够看到一个漂流物的例子，它是一个漂浮在海平面上的一个篮球大小的计算平台。这个浮动设备通常装有一个水下踏板，它随着海洋水流流动而移动。上面粘有感知板，用来检测环境现象，例如石油泄漏或者海洋微生物。作为水下通信的一个特例，这些小尺寸的感知节点能够利用低能量的声音信号进行彼此通信。

参考文献 [32] 介绍了一种不同的感知模型，它利用非协调移动感知节点。在这个模型中，不会花费能量来协调节点的运动；移动节点依附于一些传输设备，例如交通工具或者人来携带感知节点，到处移动。传感器节点进行后台管理，不时地收集数据，然而移动设备或者人沿着它们的主观路径移动。非受控移动节点仅仅测量它们路径附近的数据，节点遇到彼此时可以进行数据聚集来覆盖一个更大的范围。这种信息交换是随机的，但是这个模型能够利用一个数据交换基础设施扩展来减轻有效信息共享的困难。考虑大量的公众认可的数据融合中心，它们大多位于大量来自不同区域的人流相遇的位置（例如商业区中的机场、火车站以及大众聚会场所）。与采用大量的静态传感器节点相反，这种非受控移动模型对于长期监测是更加经济有效的，是那种最初由于物理区域太大而不能够完全覆盖的监视或者检测的一个实用的候选方案。很明显，这是以顺序完成测量所需要的延迟为代价的。

14.2.2　受控移动

在传感器网络管理中增加时空环境认知是一个有挑战性的方向。在受控移动中，网络管理者（或者应用本身）能够选择一个适当的策略来引导移动节点满足应用需求。正如先前所提到的，移动性能够满足 WSN 场景中的几个不同的目标。主要的问题的两方面是：为什么我们想要移动节点和我们如何实现节点的移动。下面的例子试图强调关于为什么要移动节点的这一部分的一些思想。

在参考文献 [26] 中，作者试图最优化地引导传感器的子集向所谓的关键物理区域的内部移动来保证数据读取的质量。每个传感器节点被认为是能够接受远程控制移动的。例如，传感器节点部署于整个森林来监测火灾的发生。传感器节点受控移动保证了健壮性，因为在关键区域中有节点失效时，其他节点是随时可用的。另一方面，不允许环境的其余部分是完全的无人监护状态也是非常重要的。一个解决方案是当移动传感器节点向人们感兴趣的区域移动时，它们也会检测它们本地邻居的覆盖范围。然而，只要有足够的剩余传感器节点使得该区域不处于无人监护状态，它们就会一直移动。

参考文献 [8] 的作者描述了有效的引导群问题，群是由拥有比普通传感器节点还要高的处理能力的节点组成，它们朝着“热”静态传感器节点移动。他们提议通过部署有限的移动群来改进传感器网络。一个群是指一组传感器节点，它们是

物理彼此邻近的，并且共享相同的移动模式。这些特定的节点可以以一个相对高的速度移动。例如，一个移动群可以是一起移动的战场上的一群坦克或者无人机。群能够直达热点来提供目的区域的细节信息。不同的群能够彼此通信或者与命令中心通话（细节详见参考文献［8］）。

在参考文献［27］中，作者提出了一个完整的移动监视和无线感知（iMouse）系统。iMouse 系统由大量廉价的静态无线传感器节点和少量比较昂贵的移动传感器节点组成。静态传感器节点进行环境监测，而移动节点能够移动到特定区域（例如潜在的紧急事件站点）并采取进一步行动（例如紧急事件拍照和进行全面分析）。

在参考文献［14］中，一个分布式无线传感器网络用于多目标监测和跟踪。当检测到一个目标，监视传感器保持静止或者开始跟踪它们的目标。是仍然静止还是跟踪一个目标是根据一个与目标和传感器节点之间的协调机制相关的优先权模式来决定的。

14.2.3　移动控制策略

在某种程度上，无线传感器网络的移动性是保持应对网络运行物理环境改变的能力和机会的一种工具。因为在某种程度上环境的改变是不可预测的，大部分情况下是随机的，因此，很难或者根本不可能预先设计良好的移动控制策略。然而，仍然需要预先定义不同的移动控制策略，根据给定应用和环境的需要在它们中选择一个合适的。

理论上，移动控制是完全集中式或完全分布式的，或者采用折中的方法。在第一类的终端有一个中央控制器能够完全控制传感器节点，也包括引导节点移动，然而另外一种相反的方式是，每个单一的移动节点需要自己决定它何时移动以及移向何处。如果我们为一个已经部署好的 WSN 寻找一个可行的解决方案，最可能正确的选择是采用分布式的策略，传感器节点之间建立通信，并利用来自邻居节点的可用信息保证健壮性来进行本地决策。下列移动控制策略一方面关注提高覆盖范围，另一方面集中环境中发生的事件。

可能最简单的基于事件的移动控制策略是单纯的响应策略（参考文献［3］是一个例子）。它仅仅使用传感器节点的当前位置和一个事件的位置来决定传感器节点的动作。每个传感器节点朝每个事件移近一小步。自由参数，即步长，也是这个策略中的一个关键因素。一次移动一个小的常数距离会导致围绕事件的聚集，或者围绕所有事件均值的聚集，这依赖于事件发生和持续的时间范围。一个更加精确的策略是定义移动传感器节点和事件之间的距离 d 的一个函数 $f(d)$，移动传感器节点根据 f 得到的数值来移向事件。当 f 构造地恰当时，可以确定如果事件本身处于多个簇中时（例如 $f(\infty)=0$），传感器节点将分散在不同簇中，并且当对事件作出响应时，传感器节点不需要遇见彼此（也就是说，$d-f(d)$ 是单调的）。

当与单纯反应解决策略比较时，基于历史的控制策略不仅利用实际事件的位置也会利用先前事件的历史和分布。每个传感器节点维护一个粗糙的直方图作为事件位置分布的近似值（详细的算法见参考文献［3］）。既然事件的数量与环境中每片区域中传感器的数量应该是成比例的，恰当地缩减事件分布，可以计算出每个区域需要的传感器节点的数量。假设一个初始是统一分布的，每个传感器能够不需要与其他节点通信，只需要根据它们的初始位置确定它们的角色。

注意到，当传感器节点开始以一种不协调的方式移向发生的事件时，由于网络可能失去整体连通性，可能产生不良的副作用。因此，实际中的一个重要要求首先是维护连通性，只要不危及连通性尽可能近地跟踪事件。保持连通性的重要性是双方面的。第一，网络连通性可靠地将信息中继到 sink 节点所必需的。第二，保持连通性也能保证覆盖范围，这意味着所有区域都不是无人监护的，因此，在监测区域中任务位置发生的未来事件都会被监测到。在这里，我们应该注意到，无线通信覆盖半径与感知覆盖半径不是一定完全相同。它们之间的关系主要依赖于给定节点的无线通信半径与感知半径的比率。因此，在通信感知中，一个连接良好的网络不必要保证网络安全覆盖，一个安全覆盖的区域也不能保证完全的无线连通性。

假设每个传感器节点有一个受限的感知半径，r_s，环境中的每个点至少可以被一个传感器节点所感知，一个简单的策略，叫做完全 Voronoi 协议，可以用来保证覆盖范围。潜在的覆盖漏洞的存在可以利用传感器节点位置（说明例子见图 14-1）的 Voronoi 图来检查。每个传感器节点使用它们自身的 Voronoi 区域几何来决定它们是否是需要覆盖。如果传感器节点 Voronoi 区域的任何一部分都比 r_s 远（注意到仅

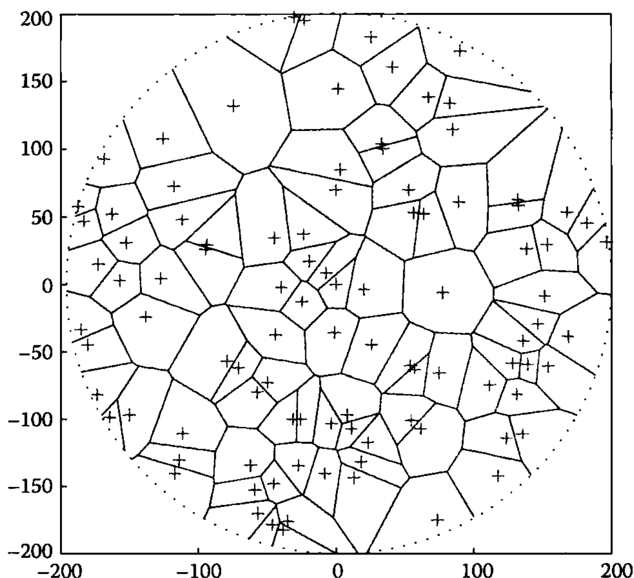


图 14-1　一个感知区域的 Voronoi 图

需检查顶点），它知道自己离那个节点最近。只要每个传感器维护它的 Voronoi 区域的覆盖范围，就能够保证网络覆盖范围。完全 Voronoi 协议的问题是它很可能计算出来是不可行的，因为计算特定节点自身的 Voronoi 区域时需要知道它的所有邻居传感器节点的位置。如果所有传感器节点都是移动的，那么位置估计将会更复杂。一个解决策略是节点移动时彼此交换新的位置信息，但是，这将会增加能量消耗，更别提定位的困难。

一个所谓的空间填充覆盖方法（Space-Filling Coverage Method）可以利用潜在的基于场所的算法来实现，该算法中每个移动传感器节点可以远离它们的邻居在空间中产生一个规则的模式。在潜在的基于场所的算法中，节点被看做承受虚拟力的虚拟粒子。典型的，节点之间因为虚拟力互相排斥（以及来自障碍物的排斥），同时也保证了节点扩散来扩展覆盖范围。

14.3　sink 节点的移动

在传统无线传感器网络中，应用应该通常是基于以单跳或多跳方式，将感知测量传递给一个或多个静态 sink 节点，这主要依赖于网络大小和节点的无线通信范围。然而，我们很快意识到移动 sink 节点将会更加满足几种应用类型的需求，它在节省能量和延长网络寿命上都是有帮助的。让我们现在来看一些赞成 sink 节点移动的具体争论。

14.3.1　为什么要移动 sink 节点

14.3.1.1　稀疏网络的数据聚集

很多情况下覆盖和监测一个部署了连通的多跳无线传感器网络的大的区域是不可行或者不必要的。例如在环境监测应用中，任务可能是需要监测几平方公里的大面积区域，但是大部分情况下大的监测区域中仅仅一些指定的部分是用户感兴趣的。因此，将大部分节点部署到这些感知区域周围，其他部分采用部分或者全部不被覆盖，是合理的。也有情况是由于处于严酷的环境，法律限制，或者其他外部因素不允许我们在整个区域部署传感器节点。因此，由于网络的一些部分与其他部分可能是完全断开的，也就是说，它们之间可能没有无线路径，就可能不能够保证多跳数据聚集到达一个静态部署的 sink 节点。然而，已经部署的传感器节点仍然能够执行它们的监测任务，暂时存储数据直到一个移动 sink 节点经过并获取数据。图 14-2 显示了这样一个场景。在这个二维区域中节点能够执行多跳路由，即使 sink 节点不在它们旁边也会传递它们的数据。在中间的稀疏区域的节点超出了彼此的无线通信范围，它们需要存储它们的数据直到 sink 节点访问它们。

14.3.1.2　负载均衡

在一个有静态 sink 节点的网络中，不同传感器节点的负载可能是高度分布不

均的。在一个单跳网络中，远离网关节点的传感器节点将会消耗大量能量进行数据传递，导致资源迅速枯竭。在图 14-3a 中，这些节点用白色的圆圈表示，一个传感器节点拥有越多的能量，它的圈的颜色越深。与之相反，在一个多跳网络中（见图 14-3b 轮到 sink 节点的邻居迅速耗尽它们的电池，因为它们将会作为所有其他传感器节点数据包的最后一跳中继节点。如果那样的话，sink 节点可能远离网络的其他部分，整个网络瘫痪。最后，如果在网络中采用一个最短路由策略，那些沿着到达静态 sink 节点最短路径的传感器节点又将很快耗尽能量。

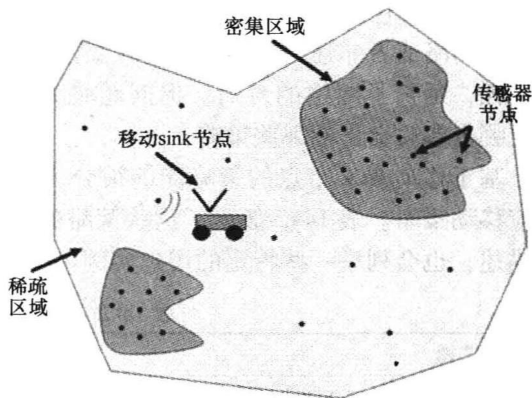


图 14-2　在一个非均匀部署网络内的移动网关

使用一个移动 sink 节点能够解决许多这种问题。如果 sink 节点节点不时地移动，不同传感器节点的负载将会改变，将会建立新的最短路径，选择新的最后一跳中继节点，因此，传感器节点之间的负载分布将会得到非常大均衡（见图 14-3c）。

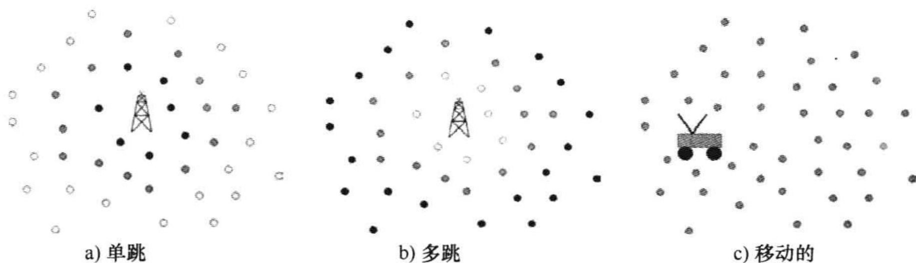


图 14-3　对于单跳和多跳网络的非均匀电池消耗。如果 sink 节点是移动的，能量消耗会更加平衡

14.3.1.3　缩短通信路径

传感器节点因区域感知，数据的数字化以及数据处理而消耗能量，但是到目前为止最耗能的任务是信息传输。在最广泛接受的能量衰减模型中[21]，信号能量衰减 $1/d^{\alpha}$，d 是与传输天线之间的距离，α 是衰减指数，一个依赖于无线通信环境的常数，它的值在 2 和 5 之间。因此，假设所有接收器对于信号检测有相同的能量阈值，特别将其标准化，支持两个相距为 d 的节点之间通信的能量需求是 d^{α} 。在这种情况下，很显然通过最小化一个传感器节点和 sink 节点之间的距离，例如，将 sink 节点移动到靠近传送数据的传感器节点，我们能够有效地降低能量消耗。这不仅对于单跳网络是可行的，而且也可用于多跳网络建立；降低多跳路径的长度会导

致更少或者更短的跳长，也就是说，将数据中继到sink节点需要更少的能量。

在传感器网络内部移动一个或者多个移动sink节点有下面许多好处：它能够从稀疏网络或者非连通网络部分中收集数据，它降低和均衡了传感器节点之间的能量消耗，延长了网络的寿命。但问题是如何移动那些sink节点？采用什么策略，以及哪个参数会影响那些策略？

基于移动sink节点的策略能够被分为四类：随机的、可预测的、受控的和自适应移动策略。表14-2显示了这些策略的主要不同。下面，这些方面将会被详细地描述，也会列举一些特定的用户实例。

表14-2 网关移动策略

移动策略	描述
随机策略	sink节点随机移动
可预知策略	sink节点以预定的速度，在一个预定义的路径上移动
受控策略	sink节点以一个预先定义的路径移动，但是速度的改变依赖于不可预测的传感器行为，有大量数据需要发送的传感器通过降低其速度来控制sink节点
自适应策略	sink节点在网络中自由移动，来适应当前事件

14.3.2 随机移动

随机移动策略很简单直接：正如它的名字所示，sink节点在网络中随机移动，例如，下述的随机航线移动模型（见图14-4）。基于这样一种场景上有许多解决策略。

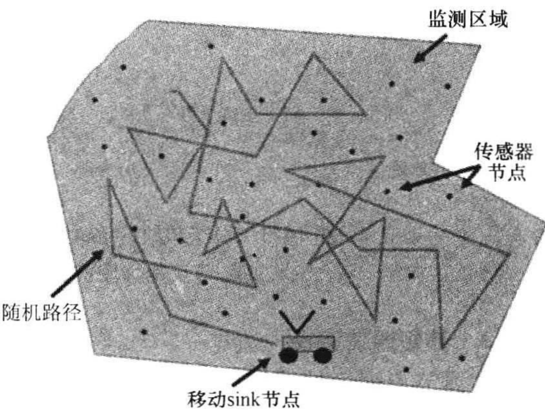


图14-4 网关沿随机路径移动模型

Shah等人使用随机移动的移动代理，叫做Data MULE（Mobile Ubiquitous LAN Extension）来保证稀疏传感器网络的数据聚集[22]。这些移动代理收集来自近范围内传感器节点的数据，缓存它们，然后将数据交给有线数据访问点。采用这些移动代理，甚至可以收集孤立的传感器节点或者区域的测量数据，孤立的传感器节点没有通往静态sink节点的无线路径。再者，没有传感器节点依赖于任何单一代理，因此任何特定的代理失效不会将传感器节点同稀疏网络断开，而仅仅是降低了工作性能。

然而这种方法的主要不足是增加了延迟，因为传感器节点传输数据前需要等待代理靠近。因此，对于一些时间紧急的应用，需要发送即时的报警信息，这种解决

策略是不可取的。例如，如果传感器网络的任务是监测一个核工厂的辐射程度，我们不能等待一个移动 sink 节点经过；无论哪个传感器节点检测到了辐射，网络需要保证可靠和实时传递一个报警信息。不过，其他应用对于延迟会不敏感很多。部署在靠近澳大利亚的昆士兰州海岸的 Great Coral Reef 的水下传感器节点通过测量海水的盐度、温度和营养级别来查找珊瑚变白的原因。在这种情况下，传递测量结果不是那么的时间紧急，延迟几个小时仍然是可以接受的。

一个相似的策略，但是是针对密集网络的，被 SENMA（Sensor Networks with Mobile Agents）所采用[25]；一个移动代理在感知区域上空随机飞行，根据评估它们与代理通信的衰退状态来发起数据收集。当移动代理传输一个信标信息时，每个传感器估计它们的衰退状态 γ。节点以概率 $s_n(\gamma)$ 成为“头”，弹出一个偏重一面的硬币，这里 n 是网络的规模。如果结果是“头”，节点传输它的数据包。否则，节点在当前时隙保持安静，在下一个时隙重新启动。注意到，尽管所有传感器节点使用相同的概率集中函数，由于衰退状态是不同的，所以节点的传输可能性是不同的。实际上，一个传感器节点仅当它有一个好的衰退状态时才会开始传输，即当代理飞近时，而当代理远离时它会自动停止传输。

也有策略使用人类的移动电话作为 sink 节点来收集来自稀疏部署传感器节点的测量数据，不能够自己进行数据传输[9]。今天的移动电话更加智能，它们能够利用许多无线技术（例如 Bluetooth、ZigBee 和 802.11）与其他邻近设备通信，并且有足够大的内存来存储数据。因此，它们能够很好地扮演移动 sink 节点的角色。这些解决策略也可以归类为随机移动，这里的随机性解释为“不受控制的”。人们将移动电话装在在口袋中，不是总是服从一个随机的移动模式；他们经常是在人行道上行走，在公路上驾驶，或者或多或少地沿着预定义的路径。然而，他们的移动由自身的兴趣和任务激发，与传感器网络的应用无关。网络不能够依赖于某个人周期性地经过，不能够预测何时或者何地一个移动电话会出现。因此，从无线传感器网络的角度来看，sink 节点的移动是“随机”的。

最后，在 SEAD（Scalable Energy efficient Asynchronous Dissemination）协议[13]中也假定了随机移动的 sink 节点，但是是在单一传感器节点传输数据到所有移动 sink 节点的场景中。这个协议的目标是建立和维护一个覆盖所有感兴趣节点的能量有效的多播分发树。

14.3.3 可预知移动

在可预知移动场景中，sink 节点沿着预定义的路径移动，通常是匀速或者至少是一个可预知的速度。当 sink 节点变得最靠近它们时，传感器节点预先知道这条路径以及一般的时间间隔。它们会试图在这些指定时间内传递它们的感知数据，以此来降低能量消耗。可预知移动对于公共交通工具是一个好的模型（公交车、航天飞机和火车），在大范围传感器网络中公共交通工具可以作为移动 sink。图 14-5

显示了这样一个场景。

最近有几篇研究文献是建立在 sink 节点可预知移动上的。在参考文献［5］中，观察者（也就是 sink）沿着预定义的路径移动，当靠近传感器节点时收集它们的数据。作者在莱斯大学（Rice University）建立了一个所提出的模型的原型，利用大学里穿梭的公交车携带移动 sink，传感器节点部署在校园的建筑上。

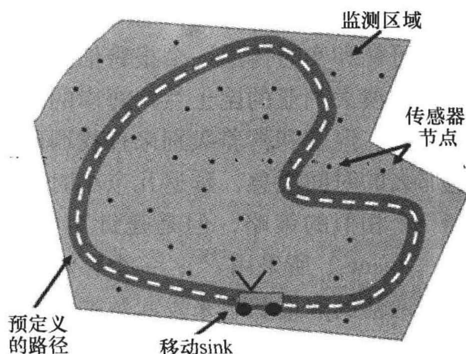


图 14-5　sink 沿着预定义的路径移动

在参考文献［17］中，假设传感器节点部署在一个圆形区域中，sink 节点预定义的路径是这片区域的外围。与先前方法相反，先前是传感器节点等待最佳时刻直接发送它们的数据到 sink 节点，这里所有的传感器节点定期发送数据到 sink 节点，即使 sink 节点在那一时刻恰好在区域的另一边；如果数据不能直接发送，会建立一个多跳路径。这里的目标不是降低通信距离，而是保证网络内部负载均衡。将这种方法的有效性与静态 sink 相比，即将这种方法的有效性与部署在圆形区域的中间的静态 sink 和随机移动的 sink 相比。结果显示随机移动 sink 将会显著降低大负载的传感器节点的能量消耗，但是在外围区域以一种可预知的方式移动 sink，更能够提高负载均衡，可以增加至多于五倍网络寿命。不足是在于传递延迟方面：只要 sink 静止的分布在区域的中间，特定传感器节点传输的数据可能需要传输两次。然而，很容易给出一个延迟的上限，因为传感器节点将会一直试图传递它们的数据，而不是等待一个随机移动的 sink 节点经过的不可预知的时间间隔。

与上面的情况相反，在参考文献［2］中，移动 sink 不是遵照一个预定义的路径。然而，所提出的方案仍然能被划分为可预知移动，因此它假设一个固有的 sink 移动模式，sink 周围的节点能够随着时间学习以及以一定的概率分布统计地描述。这个模式可能随着时间改变，但是假设每个新的模式在相当长一段时间内是有效的。然后，通过对路由决策增加一个时空维度，数据能够以一个能量有效的健壮的方式传递到移动 sink 节点。本文中所提出的路由解决策略叫做混合的加强学习时域路由（Hybrid Learning-Enforced Time Domain Routing，HLETDR）。

14.3.4　受控移动

也有基于受控移动的解决方案。在这种情况下，尽管 sink 节点仍然沿着一个预定义的路径移动，由于受到各种不能够提前预测的外部因素的影响，它的速度随时间改变。sink 节点的移动是受控的，以此来适应这些外部因素。在 AIMMS（Autonomous Intelligent Mobile Micro Server）[11,12] 中，一个移动的微服务器沿着一个特

定的路径穿过网络来路由那些深深地嵌入到网络中的节点的数据。它的移动是受控的，在有大量数据需要发送的区域，以及当前信道状态需要它多停留一段时间的区域，它需要停留较长的时间（例如它停止或者放慢速度）。然而，如果 sink 节点的移动不能够被预知，传感器节点不能够提前知道什么时候发送它们的数据；因此，当 sink 节点经过时会主动提醒它们发送数据。这很适合于查询驱动的数据聚集模型，在这种模型中，仅当 sink 节点询问时才会发送它们的数据。

在参考文献［33］中，给出了一个线性编程解决策略来决定 sink 的移动以及它在网络不同位置的逗留时间，以此来最大化网络寿命，这里网络寿命指的是到第一个传感器节点失效的时间。传感器节点放到一个小的二维格点上，以一定的速率产生数据。sink 节点仅仅能够沿着网格点移动；如果一个传感器节点既不和 sink 在同一地点也不是网格点的邻近地区，那么沿着最短路径采用多跳路由将传感器节点与 sink 节点相连。

如果 LP 问题的解决策略表明在一个给定网格点的最佳逗留时间是零，那么 sink 节点不会访问那个点。sink 节点的访问顺序不重要，因为 sink 节点在节点间的移动时间是可以忽略的，并且数据产生与时间独立。本文中提出的结果非常有趣，因为它们显示 sink 节点应该花费大部分时间在其中的一个网格角落或者中间区域。将 sink 节点放到网络的中心，我们降低了从传感器节点到 sink 的平均路由长度；然而，中间区域的传感器节点将会比靠近角落的传感器节点有更高的转发负载。因此，sink 节点应该不时移动到网格角落，并且长时间逗留来利用邻居传感器节点剩余的大量能量。遵照这些移动模式，网格中的所有传感器节点，除了那些角落中的，将会几乎同时耗尽它们的能量。这会导致网络寿命延长超过 5 倍。我们注意到，角落传感器节点的剩余能量不能够得到最大化利用：当 sink 节点移动到角落，所有通信将会被邻居节点转发，而不是由与 sink 节点在同一地点的传感器节点来转发。另一方面，当与 sink 节点距离变远时，角落传感器节点仅仅将它们自己的数据发送到它们自己的一个邻居，它们不需要执行转发任务。

然而，这种解决方法是非常受限的。对于大规模网络，线性编程不满足可扩展性，真实的网络很少会有一个网格拓扑，而且限定 sink 节点仅在网格点上移动当然也是非常不实际的。

14.3.5　自适应移动

自适应 sink 节点移动与上述的方案都不同，在自适应 sink 节点移动中，考虑到网络中当前的时间，只要关注能量消耗，sink 节点是认为持续地向着最优位置移动的。先前的解决方案描述的是时间驱动或者查询驱动的场景。在一个时间驱动的场景中，sink 节点的移动主要是承担了一个负载均衡的角色：在多跳通信中，它将会减轻那些转发来自所有其他节点的中继包的邻近节点的负担，然而在单跳通信中，它主要是减轻离 sink 节点距离较远，需要花费大量能量进行数据通信的节点

的负担。在一个查询驱动的场景中，sink 节点询问它经过的传感器节点，以脱机的方式收集它们存储的数据，这种方法仅适用于非时间紧急的应用。

与这些相反，在一个事件的驱动场景中 sink 节点的自适应移动节省能量所带来的好处远远超过了简单的负载均衡。如果在一个给定的时刻并非所有传感器节点都是活动的，而是仅有那些感知到一个指定事件的节点是活动的，将 sink 节点移向那部分节点可能是值得的。这样做，可以缩短通信路径，因此，数据聚集需要更少的能量。然而，在监测区域中如果有几个事件同时发生，如何选择 sink 节点的最优位置就不那么容易了。

在参考文献［30，31］中，作者提出了一些可取的策略来应对多跳网络中的这种场景。一个解决方案是根据当前的事件分布，最小化网络的全局能量消耗。这意味着找到一个距离事件有最小平均距离的位置，因为报告一个事件的能量需求与 sink 节点和事件之间的距离是成比例的（或者使得 sink 节点和传感器节点之间报告更精确）。图 14-6 显示了这个场景。在这个例子中，被监测区域划分为四个区域，监测到一个事件后，相邻传感器节点开始将其报告给 sink 节点。sink S1 自适应地移动来最小化到达这些事件的最小平均距离。因此，它移近那些有三个邻近事件的区域。

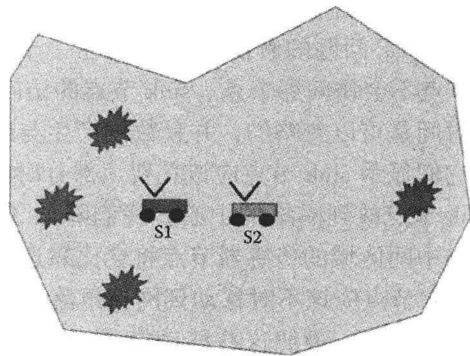


图 14-6　sink 节点位置是根据当前事件可选择的：S1 和 S2 相应地最小化时平均和最大能量消耗

这个方案的问题是，尽管全局能量消耗最小化了，传感器节点之间的能量消耗是相当不均匀的。为了避免这种情况，可以考虑最小化网络中负载大的传感器节点的传输能量。作者指出最大负载依赖于事件距离 sink 节点的最大距离。因此，这个策略与最小化事件与 sink 节点之间的最大距离是等价的，等价于最小化包围圆圈问题，也就是说，找到能够包围所有事件的最小半径的圆圈。那么最理想的 sink 节点位置是圆圈的中心。在图 14-6 中，移动 sink 节点 S2，被放在这个位置。

然而，假想 sink 节点能够直接移动到最优位置是通常不现实的；在某个时间段它只能够向最佳位置移动一步。因此，对于动态发生的事件，要做到持续地优化 sink 节点位置，最终达到优化能量消耗。

在图 14-7 中，我们能够看到一些模拟结果显示一个半径 R 为 1000m 的圆形区域的 sink 节点位置柱状图。在这块区域中事件以一个给定的概率发生，持续一个随机时间，然后消失。在底部的图中，sink 节点随机移动，独立于当前事件。与此相反，上面两个图显示的是由于当前事件的作用，根据不同的策略自适应地移动

sink 节点。如图 14-8 所示，随机移动 sink 节点导致更加同构的 sink 节点位置分布，和静止 sink 节点相比显著地延长了网络寿命。然而，随机移动不是最好的解决方案。图 14-7 中上面的柱状图显示自适应策略经常将 sink 节点部署到区域的中心，尤其是当目标是最小化最大事件距离的时候。然而，这并不会给邻居传感器节点增加过多的负载，从而使网络寿命得到进一步延长，如图 14-8 所示。因此，sink 节点随机移动仅仅保证了负载均衡，但是可能导致均衡的但高的转发负载。根据当前事件的函数自适应地移动 sink 节点会更加有效，因为它产生的是一个不均匀但是降低的转发负载，从而延长了网络寿命。

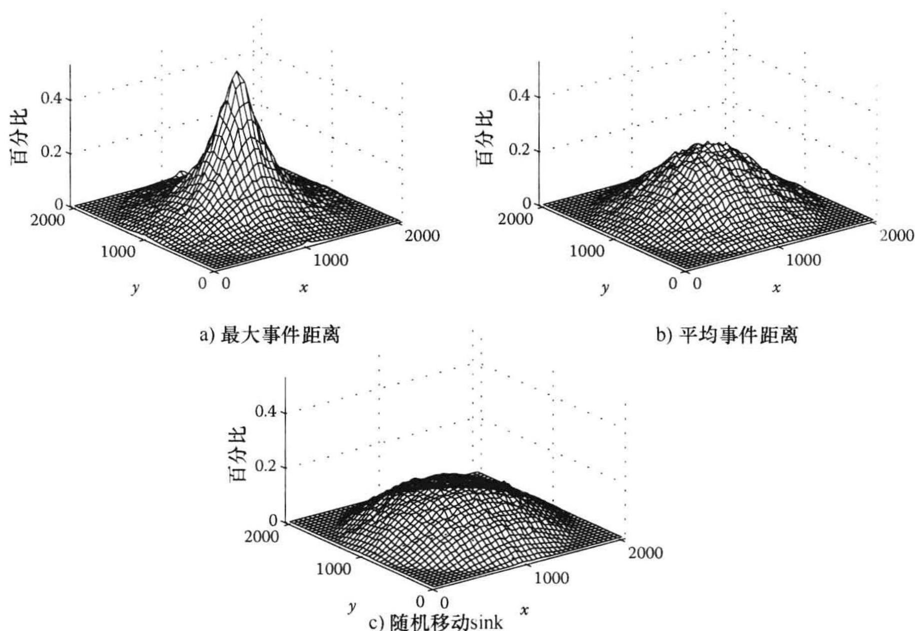


图 14-7 sink 节点位置的柱状图：最小化最大事件距离、平均事件距离或仅仅随机移动网关

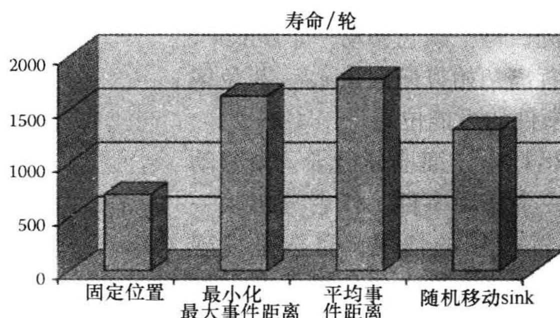


图 14-8 当与静态 sink 节点相比，不同 sink 节点移动策略的网络寿命

在参考文献［29］中，对于事件驱动的单跳网络提出同样的论证。除了上面两个策略，在一个单跳网络中，基于当前传感器节点能量级别有第三种解决方案。这种方案是为了节约那些已经报告了许多事件并且能量将要耗尽的传感器节点的能量。如果新的事件在靠近这种传感器节点的位置发生，即使在网络中有其他事件同时发生，sink节点会腾出传感器节点来靠近那个事件。图14-9显示了这样一种场景。在这个例子中，仅有两个并发事件；因此，sink节点应该位于这些事件的中心位置。然而，监测右手边事件的传感器节点（灰色显示）甚至比那些监测左手边事件的节点（黑色显示）的能量更少；因此，sink节点S_3将会解放那些传感器节点，并且靠近它们。

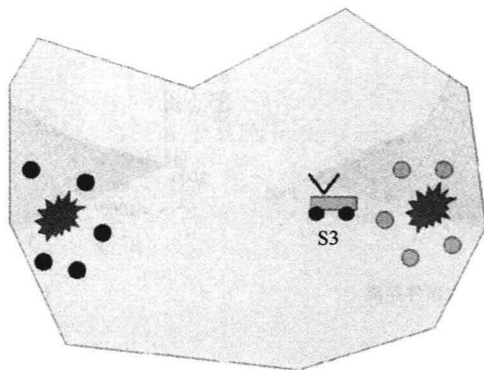


图14-9　如果将报告传感器能量级别也考虑进去，sink节点的可选择位置

如果事件不是静止的，并遵守一个特定的移动模式移动，那么情况会变得复杂。例如一个入侵检测应用，这里，事件是指一个进入受限区域，并且朝着一个特定的方向以给定的速度移动的入侵者。在这种情况下，sink节点不应该根据当前事件位置计算它的最佳位置，而是应该预测，如果可能的话，据此判断事件未来的位置和移动。图14-10显示一个遵守相关随机行走模型的事件Z的实例；每一圈它移动一个步长l，以最高的角速度σ背离初始方向向量e_θ。灰色区域显示连续时间内事件的可能位置。

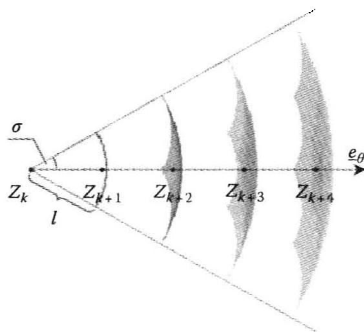


图14-10　移动事件的多步预测

在参考文献［28］中详细地分析了一个移动sink节点如何跟踪一个移动事件。作者分析了一个入侵者检测和跟踪应用的情况，目标是发现sink节点跟踪的最优轨迹，最小化传感器节点报告移动事件的能量消耗。在几个并发移动事件情况下最优化sink节点轨迹显然是一个更加困难的

任务。

表 14-3 显示本章提出的 sink 节点移动解决方案的总结，以及它们的优缺点。

表 14-3　sink 节点移动策略总结

策　略	移动性	描　述	优点和缺点
Shah 等人[22]	随机的	数据 MULE 接收和装载来自远处传感器节点的数据	支持稀疏网络，也就是说没有多跳路径存在的网络 增加了延迟 仅仅有延迟的概率上限
Tong 等人[25]	随机的	移动代理随机在区域上空飞行；当代理经过时传感器发送数据	当距离最小时进行数据发送，节约了能量 增加了延迟 仅仅有延迟的概率上限
Jayaraman 等人[9]	随机的	人们的移动手机作为 sink 节点	支持稀疏网络，即没有多跳路径存在的网络 增加了延迟 仅仅有延迟的概率上限
Chakrabarti 等人[5]	可预测的	在预定义的路径上穿梭的巴士上装有 sink 节点	距离最小时进行数据发送，节约了能量 可预测的延迟 增加了延迟
Liu 等人[17]	可预测的	环形区域，sink 节点在外围移动，数据持续发送	更低的延迟，数据周期性地发送 给出了严格的延迟上限 多跳路由上的负载均衡 长路径，更大的延迟以及更大的能量消耗
Baruah 等人[2]	可预测的	路径不是预先定义的，但是节点能够学习	高效地路由到靠近 sink 节点的节点 节点失效的健壮性 学习过程和聚合时间可能更长 路径比最佳最短路径更长
Kansal 等人[11,12]	受控的	sink 节点沿着一个预定义的轨迹移动，在有大量数据需要发送的传感器的位置或者信道质量比较差时花费更多时间	支持稀疏网络，即没有多跳路径存在的网络 更短的路径以及更低的能量消耗 给出了延迟界限 增加了延迟
Wang 等人[33]	受控的	传感器位于格上；ILP 策略决定 sink 节点在格点的最佳逗留时间	最短路径上的多跳路由，低延迟 传感器之间的负载均衡，并且延长了网络寿命 ILP 不适合于大规模网络 网格拓扑非常受限
Vincze 等人[30,31]	自适应的	sink 节点移动来适应多跳网络中的当前事件	更短的通信距离，以及更低的能量消耗 事件驱动的网络的好的解决策略，很少会被提及 不可能移动到全局最优，sink 节点仅仅一步一步接近它
Vincze 等人[29]	自适应的	在事件驱动的单跳网络中 sink 节点移近低能量供应的传感器	解放了那些能量快要耗尽的传感器 不可能移动到全局最优，sink 节点仅仅一步一步靠近它

14.4 虚拟移动

在先前的章节中，从不同方面讨论了不同情况下传感器节点和sink节点的移动。这节增加一个新类型的移动，我们称其为“虚拟的”；这里指的是没有物理移动，但是在网络中可以观察到以某种方式形成的相似的移动，对网络利用和协议有同样的影响。

一个例子是当sink节点位置改变时。正如先前讨论的，因为各种原因改变sink节点的位置可能是有利的，有时甚至对于避免传感器节点由于电池枯竭带来的严重失效是必需的。一个简单的方法是不要求在物理上移动sink节点，而是利用改变sink节点的扮演者来改变sink节点的位置，将sink节点改为当前具有有利特性的节点。从网络观点来看，虚拟地改变sink节点的位置会导致同样的协议参数变化（例如，路由表、下一跳邻居以及梯度）就好像是sink节点在物理上移动到了新的位置。改变簇头也会导致重新分簇以及簇内部通信模式的重建。

如果静态传感器节点遵守一个特定的从网络来看是有益的逻辑加入或者离开网络，传感器节点移动也是可以被虚拟化的。在拓扑控制范围内使用一些策略打开或者关闭传感器节点。这种策略的主要思想是当网络运行不需要一些节点时将它们设置为睡眠模式。然而，当环境改变，活动节点开始耗尽它们的电池，或者当节点失效时，这些备用的节点以后应该是十分方便可用的。参考文献［4］的作者提出了一个自适应自配置的WSN拓扑，基于网络状况传感器节点可以选择是否加入网络。这里传感器节点不移动，但是当节点移入网络改变网络的拓扑时，网络的全局结构需要以一个相似的方式来适应环境。

另外一个具有挑战性的想法是在网络中使用移动代码和移动代理，它由一个软件程序组成，这个软件程序通过网络在目的地执行从一个实体到另一个实体的传输（详情见参考文献［19］）。这种解决方案建立在许多WSN应用领域的一个基本性质上，也就是说，一个传感器节点的物理位置可能是对于网络运行最重要的，整个网络以一个位置感知的方式组织。在这种情况下，特别是感知和计算任务在监测到事件或者现象的位置进行本地执行。换句话说，一个移动代理可以简单地看做是传感器网络的一个有效的程序设计策略，因为感知任务可以被指定为装载有相关数据在传感器网络中传播的移动代码脚本。在这种场景下，移动代码传播到了整个网络。

当WSN应用将一些种类的事件或者监测区域中定义良好的物理区域中的关键区域作为目标时，网络运行的位置感知特性是非常重要的。这些事件或者关键区域吸引了整个网络的注意。在一个事件驱动的场景下，靠近事件的传感器节点将会开始活动，并且开始向sink节点发送测量数据。在一个查询驱动的场景下，sink节点将查询引导向关键区域（例如利用移动代理）。被监测事件可以是任何事物，如现

实生活中的物体（例如车辆、动物或者人）或一些物理现象（例如火灾），但是如果它们移动，它们是在被监测区域中进行物理移动。跟踪这些事件也会在网络中引入一些种类的移动：需要建立从事件位置到 sink 节点的路由或者至少局部重建来跟踪关注对象的位置改变。例如，目标检测和跟踪是一个无论在军事或者家庭安全应用中都经常被引用到的 WSN 应用。

14.5 传感器或者 sink 节点移动的结果

在 WSN 中使用移动传感器节点当然会对网络运行和维护产生严重的影响。首先，移动性导致拓扑变化，可能会影响算法，并且为了正确地执行路由和寻址任务，迫使它们调整一些 MAC 层或者网络管理层的协议参数。跨层设计的概念在传感器网络设计中被广泛接受，用于设计性能和成本优化的硬件和网络协议解决方案。在所有层考虑可能出现的移动性是必需的。如果不这样做可能导致严重的网络性能恶化以及应用失败，因为静态网络的协议通常不能够应用于一个持续变化的网络拓扑中。下面的章节强调对于不同层次网络协议，移动性最重要的影响，同时也指出了应付这种移动性的可能的解决方案。

14.5.1 对于节点移动的 MAC 层解决方案

对于 MAC 协议移动性可能会表现为一个重要的问题（参见参考文献［19］）。MAC 子层的主要任务是解决节点间共享用于无线通信的无线媒体的问题。在节点移动的情况下，基于媒体预留机制的 MAC 算法可能失效，因为预留机制通常假定为静态节点。一个严格的信道预留程序是在一个 TDMA（时分多址）系统（例如，选举或者类似循环赛策略）中基于分配时间片的，不会为新出现的移动节点留下空间。一个解决方案可以是周期性地重新分配时间片；这样一来，新到来的节点可以在下一轮访问媒体。同时，从上一次分配后离开的节点的未使用的时间片会被释放。

当某种类型的分簇策略引入到网络中时，移动性也会导致复杂性增大。在网络中分簇在很多方面是有益的，广泛应用于传感器网络中。它的主要思想是区分簇内和簇间通信，簇内通信更容易管理，并提供了一个数据聚集的有效方式，因此降低了网络通信量。典型的基于分簇的协议在两个交互的阶段运行，也就是说，簇形成作为建立阶段，通信阶段作为正常运行阶段。当只考虑静态节点时，环境不会发生明显的变化，理论上，在网络刚开始运行时就可以建立一次簇，然后在整个运行阶段利用已经建立起来的结构就可以了。然而，一个静态的簇层次不支持新到达的节点加入簇所造成的移动性。尤其是当簇间通信采用的是前面提及的严格的信道划分时。然而，幸运的是从移动节点来看，簇形成对于整个网络运行生命周期来说很难是静态的（甚至可以说绝不是静态的）。由于几个原因（例如，避免早期簇头资源

枯竭和未预料的节点失效）簇是不时地进行重建的，因此，给予那些物理位置移动的节点一个能够改变它们分簇的机会。确定的是，在高度移动的情况下，频繁地重建簇会引起网络严重的时间和能量花费。在某些情况下，仅要求簇头节点必须是可靠的静态节点。这样做，即使每个簇中的节点数量动态改变，簇结构还是比较容易维护的。

从移动性方面来看，最好是不要建立一个分簇结构也不要采用严格的信道访问策略。一个分布式概率 MAC 协议，利用本地可用信息，准确无误地处理移动性，反而是更加有益的。（注意，如果我们将 WSN 应用自身的需求考虑进去，并不总是这样的）当所有节点监听共享无线信道来感知它是否是空闲的，一个策略可以是基于基本的 CSMA（载波监听多路访问）方案。如果一个节点发现信道是空闲的，它发起它的无线通信。然而，这种简单的 CSMA 策略不能够完全避免碰撞（例如，可能产生所谓的隐藏终端问题）。一个有效地进一步降低碰撞的可行性方案是用 RTS/CTS（请求发送/确认发送）在任务数据传输发生前握手。这种策略在移动性上的问题是，它可能失效，因为握手后节点能够移出彼此的覆盖范围，或者在不知道任何媒体预留的情况下，外部节点能够进入竞争区域并发起通信。如果可以以某种方式区分移动和数据传输，给定的节点保持静静地移动，但是停下来的时候完全是用来数据传输的，能够进一步提高通信。

从许多提出的策略中举一个例子，移动 MAC 协议，也叫做 EAR（偷听和登记）的目标，如参考文献［23］所示，是提供移动节点要求的与静态网络交互的连通性。在这种场景下，一些随机分布的移动节点被认为是静态 WSN 的扩展。EAR 协议试图为移动节点提供持续的服务。这种策略的一个优点是将移动 MAC 协议用于后台，EAR 协议对现存的静态协议是透明的。对静态节点假设是一个类似于 TDMA 的框架结构，对移动传感器节点连接预留第一个时间片。在一些半规则时隙，一个想要邀请出现的移动节点进入网络的请求信息发送到周围邻居。

14.5.2　路由和移动性

大体上，传感器网络的大部分路由协议设计是针对固定的相同性质的传感器节点的同构传感器网络。传感器网络中有数百种路由协议（详情参见参考文献［1］）。它们中有些能够以一种令人满意的方式支持移动，其他的不可以。然而，因为移动节点与网络交互，或者无缝地整合入网络中，很可能将它们包含到网络层计算中的路由路径计算中[23]。如果移动速度是相对低的，每次一个特定的移动节点改变了它的位置可以计算一个新的路由。为了避免不必要的重复计算，可以简单地重计算移动节点邻居的路由。尽管当移动节点远离它的原始位置时这种方法开始变得低效，仅仅当必要时才需要计算一棵新的路由树。因此，在路由的建立与维护的能量消耗以及网络效率间有一个明显权衡。

正如提到的当讨论 MAC 问题时，由于在维护一个包含移动节点的分簇结构时

的代价，移动可能影响基于分簇的算法。除了簇内通信，在簇头的更高层次，产生了广阔的网络路由问题。当需要建立端到端的多跳路径时，必须有效地处理由于中继节点的突然移出导致的路由破坏，仔细地维护整个路径的有效性。特别是当一个终端节点处于所有的网络路径中，也就是，sink 节点，保持移动。最简单但是效果最差的解决方案是将所有的路由丢弃，每当 sink 节点改变它的位置时进行路由重建。然而，有更加精密的解决方案来处理移动 sink 节点——这些方法在学术界中已经引起了一定的关注。

支持移动 sink 节点的第一个路由协议是两层数据传播（Two-Tier Data Dissemination，TTDD）[16]，每个数据源建立自己的格结构。为了从一个源获取数据，sink 节点需要通过一个直接代理（Immediate Agent，IA）接着是一个位于实际格单元内的主要代理（Primary Agent，PA）去访问相应的格点之一。如果 sink 节点移走，需要根据 sink 节点移动的距离来选择一个新的 IA 和 PA。Kim 等人[13]提出一个与 TTDD 在拥有一个代理方面相似的概念，称为访问节点，用于数据转发。访问节点代表网络运行过程中的移动 sink 节点。在源节点预订了以后，数据更新沿着一个树分发到访问节点，访问节点将其转发到移动 sink 节点。仅当有访问节点发生变化时才需要对树进行更新，也就是说，当一个移动 sink 节点改变了它的访问节点的时候。两个访问节点的转换可能是由 sink 节点和它当前访问节点之间的路数的明显增长而激发。

一些其他的路由协议不需要为数据提前预约，但是会限制移动 sink 节点在一个固定的路径移动。当一个传感器节点被一个事件激发，节点自动发送相应的信息到网络中的一个移动 sink 节点。在参考文献［10，24］中，提出了一种发送感知数据沿着 sink 节点的路径到最近的节点的机制。在初始化阶段，所有的节点发现到这条路径最近的节点的最佳路由。然后，第二阶段，节点转发信息到相应的位于 sink 节点路径上的节点。当 sink 节点靠近它们时，它们将数据传递到移动 sink 节点。这种策略类似于延迟容忍网络（Delay Tolerant Network，DTN）[6]路由方法，因为这两种机制都是基于存储转发原则。

MobiRoute 策略[18]通过添加一个支持基于距离向量路由协议的移动 sink 节点扩展了 Berkeley MintRoute[34]策略。网关需要在几个锚点之间移动，并且 sink 节点会在锚点处停留一个更长的时间。sink 节点在两个锚点间的移动被分为四个状态：暂停、预移动、移动和预暂停。MintRoute 用于网络中路由。然而，在预停止状态，移动 sink 节点停止以后，路由需要迅速地更新。这些更新是通过 MobiRoute 高速率地发送路由更新信息来实现的。在参考文献［7］中，作者提出了一个简单但是有效的路由协议，利用受限的洪泛来更新到达网络中多个移动 sink 节点的路径。这种策略试图发现一个在最优路由和需要更新路由的信息数量之间的折中。

移动率是必须考虑到的一个重要的问题。如果移动率是相对低的，和高移动率相比，能够通过不同的策略来实现最优网络执行性能。考虑低移动性（例如，缓

慢变化环境状态和缓慢或者很少的网络节点移动），当状态改变时重建网络中的所有路由是可以实现的。当建立和维护对于数据通信有效的最佳的路由时，表格驱动的策略是很有优势的。为了避免数据流断裂，一个所谓的多路径方法能够帮助避免由于节点移动造成的路由断裂。如果源和目的地之间维护的是几个节点不相交的路径，那么第一个路由失效时可以转到第二个路由。对于缓慢移动的节点，甚至在路由断裂前，也可以进行本地路由修复。这可以通过一些移动预测来实现，例如，根据收到的信号强度。接收到的信号变弱表明通信距离正在增加，很快它将会导致链路断裂。有几个专门为低移动性的无线传感器网络设计的路由算法，例如GAF（Geographic Adaptive Fidelity）和TTDD（Two - Tier Data Dissemination），它们试图估计节点轨迹[19]。

当节点以高的相对速度移动时，情况会非常不同。一个例子是对于交通工具网络，尽管交通工具的移动是受限的（也就是说，它们必须位于公路上），网络需要经受非常快速的拓扑变化[19]。当拓扑频繁改变时，由于节点快速移动或者它们中的许多同时移动，都可以认为移动性是高的。在这些场景下，很难经常维护所有的路由，因为这样将会导致不可接受的路由更新负载。如果我们仍想要在高度移动环境下建立路由，可以在需要时建立而不是提前计算。对于移动自组织网络，提出了两个这样的策略并且被广泛使用，叫做AODV（Ad-Hoc On-Demand Distance Vector）和TORA（Temporally Ordered Routing Algorithm）。它们都是需求驱动的策略。这种应答策略的主要优势是避免了建立和维护一些从来未使用的路由造成的资源浪费。但是，AODV和TORA在路由发现期间都有很高的能量消耗。

移动传感器网络中最可取的路由协议可能是一种网络通信不需要使用预建立的路径的策略。实际上，这不是不可能的。有专门针对传感器网络的策略，使用flooding、gossiping或者位置感知算法，在网络中散布信息。最成功的策略是那些仅仅使用本地信息的基于分布式协议的策略。如果应用需求允许，这些类型的策略能够准确无误地支持移动性。

我们也应该注意到移动性对那些采用终端节点间直接通信的策略的影响没有那些利用多跳路径策略的影响严重。然而，假设网络中任意两个节点的直接通信（或者，对等地，任何传感器节点和一个中央sink节点直接通信）对于在无线通信时需要高能量消耗的长的通信距离，这正好是与典型的具有低能量和有限通信能力的传感器节点的假设相违背。

14.6 开放性问题

无线传感器网络的移动性引起了几个问题，从不同方面打开了研究领域。

第一，设计和制造能够移动的传感器节点，但是保持硬件自身体积小，成本低和高度的能量有效性需要一个尖端的技术背景和高超的专业技巧。随着MEMS

（microelectro mechanical systems，微机电系统）技术的不断发展和如今的 NEMS（nanoelectro mechanical systems，纳米电子系统）设备为此提供了一个可靠的背景保障。这里开放性的研究问题是技术驱动的，随着时间的发展会迅速打开较新的领域。在无线传感器网络中除了纳米技术以外，一个有趣的未来方向将会朝向纳米生物技术的使用。

第二，从网络观点来看，处理移动也是一个有挑战性的领域。所采用的协议必须一方面保持简单、有效、分布式，来满足“经典”的传感器网络的需求，但是也必须高度智能，能够自适应地应对由移动引起的复杂性。设计支持移动的自配置和自管理的无线传感器网络将会是下一个总体目标。设计隐藏潜在的拓扑改变和处理节点移动的路由策略似乎一直是一个热点问题。我们应该注意到，尽管如此，没有统一的路由策略来支持 WSN 应用领域中所有种类的移动。不同的移动策略和应用需求导致许多针对特定应用的策略，这些策略仅仅对于给定的目标是有效的。

第三，在传感器网络中移动设备的移动管理也应该引起注意。到目前为止，当移动了移动节点或者 sink 节点来最好地满足应用需求时，定义和执行哪种策略是很重要的。在有移动节点的无线传感器网络框架中，移动策略是一个需要被充分描述的重要问题[9]。移动策略的主要任务是计算移动节点的路径，可能是以分布的方式计算。特别是，移动策略和相应的映射建立算法的复杂性通常超出了传感器节点的计算能力。如果我们增加一个无线传感器网络通常会有的通信限制，问题又开始变得复杂。最近，提出了一个新的研究领域，在给定环境下考虑由一个静态的无线传感器网络提供的信息来引导一个移动机器人[19]。其思想是将传感器节点看作移动机器人感知能力的扩展（即使机器人身上没有传感器节点）。

14.7 结论

这一章完整地描述分析了在无线传感器网络中采用移动元件、移动传感器节点和 sink 节点所带来的优势。移动节点的出现能够延长网络的寿命，平衡能量消耗，减小通信路径，扩展网络覆盖范围。然而，利用移动节点也有几个后果，例如，关于路由和 MAC，需要进行仔细处理来保留取得的优势。

参考文献

1. J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: A survey. *IEEE Wireless Communications*, 11(6):6–28, 2004.
2. P. Baruah, R. Urgaonkar, and B. Krishnamachari. Learning-enforced time domain routing to mobile sinks in wireless sensor fields. In *Proc., 29th Annual IEEE International Conference on Local Computer Networks (LCN)*, pp. 525–532, Washington, DC, 2004.
3. Z. J. Butler and D. Rus. Controlling mobile sensors for monitoring events with coverage constraints, In *Proc., IEEE International Conference of Robotics and Automation*

(ICRA), pp. 1568–1573, New Orleans, LA, April 2004.

4. A. Cerpa and D. Estrin. ASCENT: Adaptive self-configuring sensor networks topologies, *IEEE Transactions on Mobile Computing*, 3(3):272–285, 2004.
5. A. Chakrabarti, A. Sabharwal, and B. Aazhang. Using predictable observer mobility for power efficient design of sensor networks. In *Proc., 2nd International Workshop on Information Processing in Sensor Networks (IPSN)*, pp. 129–145, Palo Alto, CA, April 2003. Also in *Lecture Notes in Computer Science*, 2634.
6. K. Fall, A delay-tolerant network architecture for challenged internets. In *Proc., International Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)*, pp. 27–34, New York, 2003. ACM.
7. K. Fodor and A. Vidács. Efficient routing to mobile sinks in wireless sensor networks, In *Proc., 2nd International Workshop on Performance Control in Wireless Sensor Networks (PWSN)*, Austin, TX, October 2007.
8. M. Gerla and K. Xu. Multimedia streaming in large-scale sensor networks with mobile swarms. *ACM SIGMOD Record: Special Section on Sensor Network Technology and Sensor Data Managment*, 32(4):72–76, 2003.
9. P. P. Jayaraman, A. Zaslavsky, and J. Delsing. Sensor data collection using heterogeneous mobile devices. In *Proc., IEEE International Conference on Pervasive Services*, pp. 161–164, Istanbul, Turkey, June 2007.
10. D. Jea, A. A. Somasundara, and M. B. Srivastava. Multiple controlled mobile elements (data mules) for data collection in sensor networks. In *Proc., IEEE/ACM International Conference on Distributed Computing in Sensor Systems (DCOSS)*, Marina del Rey, CA, June 2005.
11. A. Kansal, M. Rahimi, W. J. Kaiser, M. B. Srivastava, G. J. Pottie, and D. Estrin. Controlled mobility for sustainable wireless networks. In *Proc., IEEE Sensor and Ad Hoc Communications and Networks (SECON)*, Santa Clara, CA, October 2004.
12. A. Kansal, A. Somasundara, D. Jea, M. B. Srivastava, and D. Estrin. Intelligent fluid infrastructure for embedded networks. In *Proc., ACM 2nd International Conference on Mobile Systems, Applications, and Services (MOBISYS)*, pp. 111–124, Boston, MA, June 2004.
13. H. S. Kim, T. F. Abdelzaher, and W. H. Kwon. Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. In *Proc., 1st International Conference on Embedded Networked Sensor Systems (SenSys)*, pp. 193–204, New York, 2003. ACM.
14. K. M. Krishna, H. Hexmoor, P. S. Rao, and S. Chellapa. A surveillance system based on multiple mobile sensors. In *Proc., 17th International FLAIRS Conference: Special Track on AI Techniques in Multi-Sensor Fusion*, May 2004.
15. B. Liu, P. Brass, O. Dousse, P. Nain, and D. Towsley. Mobility improves coverage of sensor networks. In *Proc., 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pp. 300–308, New York, 2005. ACM.
16. H. Luo, F. Ye, J. Cheng, S. Lu, and L. Zhang. TTDD: Two-tier data dissemination in large-scale wireless sensor networks. In *Wireless Networks*, 11(1–2):161–175, 2005.
17. J. Luo and J-P. Hubaux. Joint mobility and routing for lifetime elongation in wireless sensor networks. In *Proc., IEEE INFOCOM*, pp. 1735–1746, Miami, FL, 2005.

18. J. Luo, J. Panchard, M. Piorkowski, M. Grossglauser, and J-P. Hubaux. MobiRoute: Routing towards a mobile sink for improving lifetime in sensor networks. In *Proc., 2nd IEEE/ACM International Conference on Distributed Computing in Sensor Systems (DCOSS)*, San Francisco, CA, 2006.
19. P. J. Marron. Research directions of cooperating objects with mobile nodes. In *Proc., NSF Workshop on Data Management for Mobile Sensor Networks (MobiSensors)*, Pittsburgh, PA, January 2007. Position Paper.
20. S. Nittel, N. Trigoni, and N. Pettigrew. Data management in mobile ad-hoc ocean sensor networks. In *Proc., NSF Workshop on Data Management for Mobile Sensor Networks (MobiSensors)*, Pittsburgh, PA, January 2007. Position Paper.
21. T. S. Rappaport, *Wireless Commnications: Principle and Practice*, Prentice Hall, Englewood Cliffs, NJ, 2002.
22. R. C. Shah, S. Roy, S. Jain, and W. Brunette. Data MULEs: Modeling a three-tier architecture for sparse sensor networks. In *Proc., IEEE Workshop on Sensor Network Protocols and Applications (SNPA)*, pp. 30–41, Anchorage, AK, May 2003.
23. K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie. Protocols for self-organization of a wireless sensor network. *Personal Communications, IEEE [see also IEEE Wireless Communications]*, 7(5):16–27, 2000.
24. A. A. Somasundara. Controllably mobile infrastructure for low energy embedded networks. *IEEE Transactions on Mobile Computing*, 5(8):958–973, 2006.
25. L. Tong, Q. Zhao, and S. Adireddy. Sensor networks with mobile agents. In *Proc., IEEE MILCOM*, vol. 22, pp. 688–693, Boston, MA, October 2003.
26. G. Trajcevski, P. Scheuermann, and H. Brönnimann. Mission-critical management of mobile sensors: Or, how to guide a flock of sensors. In *Proc., 1st International Workshop on Data Management for Sensor Networks (DMSN)*, pp. 111–118, New York, 2004. ACM.
27. Y-C. Tseng, Y-C. Wang, and K-Y. Cheng. An integrated mobile surveillance and wireless sensor (iMouse) system and its detection delay analysis. In *Proc., 8th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, pp. 178–181, New York, 2005.
28. A. Vidács and J. T. Virtamo. Minimum transmission energy trajectories for a linear pursuit problem. In *Proc., 1st EuroFGI International Conference on Network Control and Optimization (NET-COOP)*, pp. 286–295, 2007.
29. Z. Vincze, D. Vass, R. Vida, and A. Vidács. Adaptive sink mobility in event-driven clustered single-hop wireless sensor networks. In *Proc., 6th Int. Network Conference (INC)*, pp. 315–322, Nice, France, 2006.
30. Z. Vincze, D. Vass, R. Vida, A. Vidács, and A. Telcs. Sink mobility in event-driven multi-hop wireless sensor networks. In *Proc., 1st Int. Conf. on Integrated Internet Ad hoc and Sensor Networks (InterSense)*, page Article No. 13, Nice, France, 2006.
31. Z. Vincze, D. Vass, R. Vida, A. Vidacs, and A. Telcs. Adaptive sink mobility in event-driven densely deployed wireless sensor networks. *International Journal on Ad Hoc and Sensor Wireless Networks*, 3(2–3):255–284, 2007.
32. K-C. Wang and P. Ramanathan. Collaborative sensing using sensors of uncoordinated mobility. In *Proc., IEEE/ACM International Conference on Distributed Computing in Sensor Systems (DCOSS)*, vol. 3560 of *Lecture Notes in Computer Science*,

pp. 293–306. Springer, Marina del Rey, CA, 2005.

33. Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli. Exploiting sink mobility for maximizing sensor networks lifetime. In *Proc., 38th Hawaii International Conference on System Sciences*, Big Island, HI, January 2005.
34. A. Woo, T. Tong, and D. Culler. Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proc., 1st International Conference on Embedded Networked Sensor Systems (SenSys)*, pp. 14–27, New York, 2003. ACM.
35. V. I. Zadorozhny. Mobility-aware query optimization in data intensive mobile sensor networks. In *Proc., NSF Workshop on Data Management for Mobile Sensor Networks (MobiSensors)*, Pittsburgh, PA, January 2007. Position Paper.

第 15 章　无线传感器网络安全技术

无线传感器网络（WSN）是由散布在一大片物理区域内的微小的感知设备组成的网络，用于收集和处理环境数据，例如温度、湿度、光照情况、地震活动以及环境图像。这些数据能够用于监测特定的事件以及触发相应的活动。

由于传感器网络向广泛部署的任务关键区移近，安全问题成为一个核心的关注点。到目前为止，大量研究关注于传感器网络的可用性和有用性，还未关注其安全问题。另一方面，由于内在的资源和计算能力的限制以及缺少基础设施，传感器网络的安全与传统网络安全相比面临不同的挑战。

这些限制构成了传感器网络中采用传统计算机安全技术的主要障碍。不可靠的信道和无人操作导致安全防护更加困难。幸运的是，新问题也激发了新的研究，代表了一个从开始就恰当地描述（解决）传感器网络安全的机会。许多研究者开始应对无线传感器网络中在预防攻击者的同时最大化处理能力和能量保留的挑战。我们将要讨论的安全问题包括传统的安全问题，如安全高效的路由，数据聚合，密钥管理和入侵检测，也有那些特别针对无线传感器网络的安全问题，像建立一个感知信任模型和防范物理攻击。

本章中，我们将覆盖传感器网络安全方面的主要话题，提出这一领域的障碍和需求，对许多当前的攻击进行分类，最后列出它们相应的防范措施。

本章的结构如下。首先，简单介绍传感器网络安全问题及其背景。我们也会讨论传感器网络安全问题所面临的困难。第二，我们列出的传感器网络的安全需求。第三，我们对传感器网络的主要攻击进行分类，简要概述相应的防范措施。最后，我们通过讨论安全问题的一些可能的解决策略结束本章。

15.1　概述

传感器网络必须整理几种类型的数据包，包括路由协议包和密钥管理协议包，后者是出于对安全的考虑。一个给定的传感器网络中采用的关键建立技术应该满足几个要求才是高效的。在这些要求中可能包括支持网内处理和便利的数据自组织。然而，一个安全的应用技术必须最少包含以下安全目标：真实性、保密性、完整性、隐私保护、可扩展性以及灵活性[1]。

15.1.1　安全目标

1. 真实性

真实性允许一个节点确定与它通信的同伴节点的身份，如果没有安全性保障则一个攻击者可以模仿一个节点，因此能够未经授权地访问资源和感知信息，以及干涉其他节点的操作。安全技术应该保证在一次通信中网络中的通信节点有一种方式能够证实其他节点的身份，例如，接收节点应该能够识别发送节点的ID。

2. 保密性

保密性保证了特定的信息不会暴露给未经授权的实体。安全技术应该防止数据信息泄露给未经授权的组织。一个攻击者可能试图通过获取密钥获得数据来攻击一个传感器网络。一个更好的密钥技术控制被盗用节点，防止其数据被进一步泄露。

3. 隐私保护

隐私保护是当身份败露时主要的传感器网络安全技术之一。知道了传感器网络中节点的身份，外部组织能够发起一些严重的目标导向的攻击。因此，传感器节点的稀缺资源被很快地彻底耗尽，网络通信中断。为了保护节点的隐私，一个匿名通信协议可以是网络中的一个有效的通信工具，而不需要透露节点的ID。

4. 完整性

完整性意味着传输中不会有数据被窜改。即被传输的信息不会被毁坏。从关键的建立技术来说，意义如下。只有网络中的节点能够访问密钥，并且只有一个指定的基站拥有改变密钥的特权。这将有效地防止未经授权的节点获取已被使用的密钥的信息，而且防止了来自外部资源的更新。

5. 可扩展性

高效性要求传感器网络使用一个可扩展的安全技术，允许网络规模随特定网络而变化。采用的技术应该为小网络提供高的安全特性，并且当将其应用于更大的网络时能够保持这些特性。

6. 灵活性

安全技术应该能够在任何环境下良好地运行，并且支持节点的动态部署，也就是说，这些技术在多重应用中是可用的，并且允许随时增加节点。发展传感器网络面临的挑战之一是利用有限的资源提供高安全性。传感器网络的制造成本不能够太高，因为它们很可能会被部署到敌方环境中，可能会被敌方夺取以获得关键信息或者简单地被敌方破坏掉，这样会导致巨大的损失。这些受成本限制约束的部分包括不能够完全使得传感器网络是防干扰的。在开发一个包括电池寿命、通信范围、带宽、内存和先前部署知识的安全技术的同时，要记得其他传感器节点的限制。

7. 电池寿命

传感器节点的电池寿命有限，使得采用非对称密钥技术，如公钥密码体制，是不实际的，因为它们使用更多的能量进行整体的复杂的数学计算。可以利用更有效

的对称技术来缓解这个限制，它包括更少的计算过程和需要更少的运行能量。

8. 传输半径

有限的能量供应也限制了传输半径。传感器节点仅能够将信息传输到指定的短半径内，因为半径增加可能导致能量耗尽。一些技术像网内处理技术能够通过聚集和仅传输一些节点处理过的信息来帮助获得更好的表现。通过这种方式可以防止能量浪费。

9. 带宽

用一般的传感器节点有限的带宽能力传输大型的数据块效率不高，例如 UC Berkeley Mica 平台的发送器带宽只有 10kbit/s。为了补充这个不足，一个安全协议应该只能允许小数据块同时传递。

10. 内存

传感器节点的可用内存通常是 6 ~ 8kbit/s，其中一半被典型的传感器网络操作系统占用，像 TinyOS。安全协议必须有效地利用剩余的有限内存空间将密钥、缓冲存储信息等存储在内存中。

11. 可用性

可用性确保存在拒绝服务攻击（DoS）时的残存性。在物理媒体访问控制层中，攻击者能够使用拥塞技术干扰物理信道通信。在网络层，攻击者能够扰乱路由协议。在更高层，攻击者能够降低高层服务，例如密钥管理服务。

12. 不可抵赖性

不可抵赖性确保了信息源不可否认自己发送了信息。

13. 先前的部署知识

由于传感器网络中的节点动态、随机部署，不可能维持每个位置的信息。因此，一个安全协议在初始化网络的密钥时不应该知道节点部署的位置。一个协议是否安全，不仅看它提供的对所传输信息保密的能力，还必须根据攻击者的弱点满足特定的其他标准，包括传感器网络的三个方面：抵抗力、撤销和弹性。

14. 抵抗力

敌方可能攻击网络，撤销网络中的一些节点，然后复制那些节点放回网络中。利用这种攻击，敌方将它的复制节点填入整个网络，从而获得整个网络的控制权。一个安全协议必须能够抵抗节点的复制来防止这种攻击。

15. 撤销

如是说一个传感器网络被敌方侵入了，安全协议应该提供一个有效的方式来撤销被盗用的节点，使用一种轻量级的方法，即不会使用大部分网络已经有限的通信能力。

16. 弹性

如果传感器网络中的一个节点被捕获了，安全协议应该保证其他节点的保密信息不被泄露。一个模式的弹性是利用网络中全部的被盗用的节点以及全部被盗用的

通信部分来计算的。弹性也意味着方便地使新加入的传感器参与安全通信。

15.1.2 挑战

使用无线连接使得传感器网络受到一系列攻击的影响，如被动窃听和主动冒充，消息重放以及消息变形。窃听可能为攻击者提供访问机密信息的通道，因此违背了保密性。主动攻击包括删除消息、注入错误消息、扮演一个节点等，这样违背了可用性、完整性、真实性以及不可抵赖性。节点自由地在敌方环境中移动，它们有相对稀少的物理保护，很可能被盗用。因此，不仅需要考虑来自外部的恶意攻击，而且需要考虑来自网络内部的被盗用节点的攻击。为了达到高的残存性，传感器网络应该具有一个无中心实体的分布式结构；集中性提高了遭受攻击的可能性。由于传感器网络频繁的拓扑改变，它可以是动态的，例如，一个自组织的传感器车载网络。甚至单独节点之间的信任关系也会改变，尤其是当一些节点被发现是被盗用的节点时。安全机制必须是动态的，而不是静态的，并且对数百或数千的节点应该是可扩展的。

15.1.3 密钥管理

加密模式，例如数字签名，经常被用来保护路由信息以及数据。因为密钥分发的上传方式，公钥密码体制通常是受欢迎的。在公钥基础设施中，每个节点有一个公钥/私钥对。公钥被分发给其他节点，同时私钥保留在节点自身，这样保密性更好。第三部分（可信任的）叫做认证中心（Certification Authority，CA）是用来管理密钥的。CA 有一个公钥/私钥对，公钥被所有节点所知，标记有将公钥绑定到节点的证明。受信任的 CA 需要保持在线来响应当前的绑定，因为超时会进行绑定更新。如果拥有公钥的节点不再受信任或者已经不在网络中，那么它的公钥应该被取消。传感器网络采用一个单一的密钥管理服务可能不是一个好的方法，因为它很可能成为网络中的弱点。如果一个 CA 是离线的或者是不可用的，那么节点就不能得到当前的其他节点的公钥来建立一个安全的连接。同样，如果一个 CA 被盗用了，攻击者就能够对私钥标记任何错误的认证。简单复制一个 CA 将会使网络更加脆弱，因为盗用一个单一的复制的 CA 可能导致系统失效。因此，通过让这些节点共享密钥管理的责任，将信任分发到一系列节点需要更加小心谨慎。

15.1.4 安全路由

当前的路由协议能够很好地应对信息转发和处理的工作，但是设计中未考虑防御恶意攻击者。没有单独的标准协议来捕获普通的安全威胁，以及提供路由协议的安全准则。节点简略地改变网络的控制信息来建立节点之间的路由；这是恶意攻击者破坏网络的潜在的目标之一。外部攻击者通过向网络中注入错误的路由信息，重放旧的路由信息，或者对路由信息进行变形来分割网络，或者利用重放路由信息或

者发送无效路由信息使网络过载来执行攻击。内部被盗用的节点更难被察觉和纠正。因为被盗用的节点将会利用它们的私钥产生无效的签名，每个节点标识的路由信息都将会失效。由于在一些特定传感器网络应用中拓扑结构动态变化，通过路由信息来发现被盗用的节点也是很困难的。路由协议必须处理过期的路由信息来适应动态改变的拓扑。被盗用节点产生的错误的路由信息也可以被看做是过期的路由信息。只要有足够数量的有效节点，路由协议应该能够绕过被盗用的节点，然而这样需要节点之间存在多重的，可能会断裂的路由。如果当前的路由失效一个路由协议应该能够利用另外一个路由。

15.2　预备知识

对于小的传感器节点来说公钥加密机制代价太高了，因为传统的公钥算法［例如 Rivest Shamir Adleman algorithm（RSA）］需要进行大量计算，这并不适合微小节点。然而，近来在椭圆曲线加密算法（Elliptic-Curve Cryptography，ECC）[2]上的进展为传感器网络中采用公钥加密机制提供了新的机会。最近实现的在 Atmel ATmega128，8Hz 和 8bit 的 CPU 上的 160 bit ECC，显示一个 ECC 点乘法仅需要花费低于 1s 的时间[3]，这表明 ECC 公钥加密对传感器网络是可用的。

ECC 能够与 Diffie-Hellman 方法结合起来，为两个通信方提供密钥交换模式。ECC 也可以用来产生数字签名。椭圆曲线数位签章演算法（Elliptic Curve Digital Signature，ECDSA）利用 ECC 来产生用于认证的数字签名以及其他的安全用途[4,5]。已经提出了几种利用 ECC 的加密和解密的方法[2,4]。详情请见参考文献[2,4,5]。根据 ECC，在参考文献［6］中，Duet 等人为异构传感器网络提出了一个密钥管理模式，将它与现存的模式比较，具有更好的表现。这个协议将会在 15.4.1.3 节中介绍。另外一个协议是基于身份的加密（Identity-Based Encryption，IBE）和配对加密（Pairing-Based Cryptography，PBC），在 15.4.1.4 中讨论，它们比先前提出的算法具有更好的表现。

在密码学分支中，基于身份的加密（Identity-Based Cryptography，IBC）[7]是一个例外，一个唯一标记用户的信息（例如 IP 或者电子邮箱地址）能够用于交换密钥和加密数据，因此 PKI 不是必需的。仅当 PBC[8]出现时它才是实际可用的。双线性对，例如椭圆或者超椭圆曲线上的韦伊配对或者塔特配对最近被发现用于密码协议的设计中。最先知道实行配对的传感器节点是基于 8bit/7.3828MHz ATmega128L 微控制器的（例如 MICA2 和 MICAz 智能尘埃），在参考文献［9］中进行了研究，得出的结论是来自配对的加密在资源受限的节点上是可行的。

在参考文献［10，11］中，Leonardo 等人主张 IBE 是 WSN 理想的加密模式，因为 WSN 满足了 IBE 模式的严格要求。他们进一步认为 WSN 也是使用 IBE 的理想场景。他们也讨论了在资源受限的节点上利用和实施 IBE，提出了一些结论。特别

是，他们在 MICAz——新一代 MICA 节点上评估了 IBE 最显著的操作——配对[12]。评估 IBE 时最耗时的部分是配对操作。他们描述了实施问题和提出了在 MICAz 上计算配对的结果。MICAz 由 ATmega128 微控制器启动（8bit/7.38MHz 处理器，4KB SRAM，128KB 闪存）。配对在运行 TinyOS 的 MICAz 节点上进行衡量[13]。计算一个配对的平均执行时间是 30.21s。关于随机访问存储器（RAM）和只读存储器（ROM）的代价是 1831 和 18384B。因此，上面的结果显示配对计算对传感器网络节点是可行的。

15.2.1 椭圆曲线

椭圆曲线引起人们的兴趣主要是因为作为一个两者择一的组结构。当说到通常的加密协议，椭圆曲线家族会带来特定的优势，$E(F_q): y^2 = x^3 + Ax + B$。它的主要优势是，可以使用更小的密钥，因为对于大多数这样的曲线的分离对数（Discrete Logarithm，DL）问题还没有已知的多项式时间的算法。定义在有限区域 F_q 的给定一个点 P 和一个曲线 E，$q = p^m$，并且 p 是一个大素数；问题是为给定的“aP”确定“a”。在大部分环境下，在这样一个曲线上的点形成一个简单的循环群。曲线上的每个点有一个顺序。这个顺序是最小的正整数，r，$rP = O$，O 是群的身份点，即所谓的在无穷远处的点。曲线上的点的数量，曲线的顺序，叫做$\#E$。每个有效的 r 划分$\#E$。我们也需要知道一个重要的关系式$\#E = q + 1 - t$，t 是“弗罗贝纽斯（Frobenius）轨迹”，t 是相对小的——每个曲线拥有的一个常数。我们也注意到“扭曲的”曲线，$E^t(Fq): y^2 = x^3 + d^2Ax + d^3B$，$d$ 是模 q 的二次非剩余。这个曲线上有$\#E = q + 1 + t$ 个点[14]。

15.2.2 椭圆曲线群和分离对数问题

每个原型系统的基础是一个困难的不能够通过计算来解决的数学问题。DL 问题是许多原型系统安全的基础，包括椭圆曲线原型系统。更具体地说，ECC 依赖于困难的椭圆曲线分离对数问题（Elliptic Curve Discrete Logarithm Problem，EC-DLP）[15]。

记得我们研究了两个特定椭圆曲线组上定义的几何级数的操作。这两个操作是点加和点倍的。通过选择一个椭圆曲线组上的一个点，对它加倍来获得点 $2P$。然后，可以将点 P 加到 $2P$ 上来获得点 $3P$。用这种方式确定的点 nP 被认为是一个点的标量倍增。ECDLP 是基于标题倍量乘积的难解性。

尽管习惯上使用添加标记来描述一个椭圆曲线组，通过利用倍增标记可以提供一些更深入的了解。尤其是，考虑到添加标记下的称为“标量倍增”的操作，就是，将 k 个点 P 的复本加起来计算 kP。倍增标记操作由点 P 的 k 个复本相乘来构成，产生了点 $P * P * P * P......$，也就是说，k 次 $P = Pk$。

在倍增组 $Zp*$ 中，DL 问题是：给定这一组中的元素 r 和 q，和一个质数 p，找

到一个数字 k，使得 $r = qk \bmod p$。如果椭圆曲线组是使用倍增标记来描述的，那么椭圆曲线的 DL 问题是：在组中给定点 P 和 Q，找到一个数字使得 $Pk = Q$；k 被称为对于基 P 的 Q 的 DL。当椭圆曲线组利用添加标记来描述时，ECDLP 是：给定组中的 P 和 Q，找到一个数字 k 使得 $Pk = Q$。在一个真实的应用中，k 将足够大使得当采用这种方式来决定 k 时是不可行的。

尽管大部分基于 DL 的协议原本定义在一个有限的领域 F_q^* 的倍增组，DL 问题（和 Diffie-Hellman 问题）当然也可以定义在任何组上。协议因此能够按照组来解释可能允许更安全或者更有效的算法。然而，在一些组中采用 DL 是容易的，例如在一个有限区域的附加组中。显然，这样的组由于安全是基于 DL 的难度，不适合加密的目标。幸运的是，也存在解决 DL 问题比在 F_q^* 中更难的组，索引微积分方法提供了一个亚指数的算法。一个椭圆曲线上的一组点是这样一个组的一个例子。

15.2.3　双线性配对

让 G_1 做为一个附加组，G_2 作为一个相同主顺序的倍乘组 q。让 P 作为 G_1 的一个随机发生器（“aP 表示 P 自身相加 a 次”）。假设在 G_1 和 G_2 中 DL 问题都是困难的。我们认为 G_1 作为一个椭圆曲线上的 F_q 上的一组点，G_2 作为一个有限区域 F_{q^k} 上的倍增组的子组，$k \in Z_q^*$。一个映射 $\tilde{e}$：$G_1 \times G_1 \rightarrow G_2$，满足下面的特性，被称为一个密码的双线性映射。

双线性：对于所有的 $P,Q \in G_1$ 和 $a,b \in Z_q^*$ 满足 $\tilde{e}(aP,bQ) = \tilde{e}(P,Q)^{ab}$。这也可以用下面的方式描述为：对于 $P,Q,R \in G_1$，$\tilde{e}(P+Q,R) = \tilde{e}(P,R)\tilde{e}(Q,R)$，并且 $\tilde{e}(P,Q+R) = \tilde{e}(P,Q)\tilde{e}(P,R)$。

非退化：如果 P 是 G_1 的一个发生器，然后 $\tilde{e}(P,P)$ 是 G_2 的一个发生器。换句话说，$\tilde{e}(P,P) \neq 1$。

可计算性：如果对于所有的 $P,Q \in G_1$，$\tilde{e}(P,P)$ 在多项式时间内都是可计算的，那么这个映射是有效的可计算的。

修改的 Weil Pairing[16] 和 Tate Pairing[17,18] 是加密双线性映射的例子。

15.2.4　Diffie-Hellman 问题

本章中，我们回想 Diffie-Hellman 间隙群的特性。首先，我们假设用下面的术语来定义 Diffie-Hellman 间隙群。

让 P 作为一个椭圆曲线 E 上的一个点，由 $y^2 = X^3 + \alpha X + \beta \bmod T$ 给定，T 是一个质数。

$\langle P \rangle$ 是由 P 产生的一个 E 的子组。

$|\langle P \rangle| = q$

$a,b \in Z_q^*$

因此，我们能够认为 G_1 是在椭圆曲线 E 上的一组点。利用这组点，我们能够定义下面的困难的加密问题。

计算 Diffie-Hellman（CDH）问题：给定一个三元级$(P,aP,bP) \in G_1$，对于 a，$b \in Z_p^*$，找到元素 abP。

决定 Diffie-Hellman（DDH）问题：给定一个四元组$(P,aP,bP,cP) \in G_1$，对于 a，b，$c \in Z_p^*$，确定 $c = ab \bmod q$ 是否成立。

间隙 Diffie-Hellman（GDH）问题：这是一系列关于 CDH 问题是困难的和 DDH 是简单的问题。

双线性 Diffie-Hellman（BDH）问题：给定一个四元组$(P,aP,bP,cP) \in G_1$，对于一些 a，b，$c \in Z_p^*$，计算 $\tilde{e}(P,P)^{abc}$。

那些 CDH 问题困难而 DDH 问题是容易的群称为 GDH 群。GDH 组的细节能够在参考文献［19-21］中找到。也有一些其他的与提及的协议无关的 Diffie-Hellman 问题。关于其他 Diffie-Hellman 问题的细节能够在参考文献［22］中找到。

15.3 攻击类型

本章描述了传感器网络中有效的不同种类的攻击。

15.3.1 被动攻击

被动攻击主要包括未授权的路由包“偷听”或者安静地拒绝执行请求的功能。这种类型的攻击可能是试图获得路由信息，攻击者利用这些路由信息可以预测每个节点相对于其他节点的位置数据。这样的攻击通常不可能被发现，因为攻击者并未扰乱路由协议的操作，仅仅试图通过偷听路由传输来发现有用信息。

15.3.2 主动攻击

主动攻击意味着降低或者阻止节点间的信息传输。它们能够导致节点间通信恶化甚至完全终止。通常，这样的攻击包括敌方执行的行为，例如重放攻击、篡改和删除交换数据。

15.3.3 拒绝服务攻击

当攻击者利用无用传输使节点过载，合法请求无法得到处理，无法访问资源时，就意味着发生了拒绝服务攻击。发送到目标节点的包将会随机选择返回地址，经常被伪造源地址，因此目标节点很难发现精确的攻击者的位置。

多个攻击者，例如 I_1、I_2、I_3 以及 I_4 相互合作，能够设定一个指定的节点 N，作为一个攻击目标来耗尽它的资源，如图 15-1 所示。一个有足够的能量的攻击者

I，如图 15-2 所示，能够设置指定的节点 N_1、N_2、N_3 和 N_4 为目标来耗尽这些节点的资源。主要的概念是识别这个节点和为指定的节点选择一个目标。

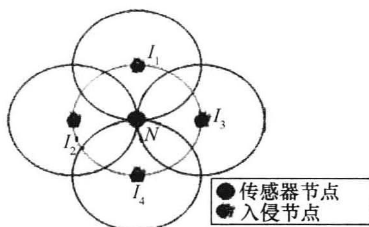


图 15-1　DoS，根据目标；多对一

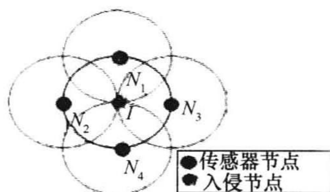


图 15-2　DoS，根据目标；一对多

15.3.4　虫孔攻击

在虫孔攻击中，一个攻击者记录网络一个位置的一个包，通过高质量的带外链接将这个包发送到另外一个位置，这个带外链接称为一个隧道[23]，建立在网络中的攻击者之间。图 15-3 显示一个基本的虫孔攻击。攻击者在节点 I_2 上重放节点 I_1 收到的包，反之亦然。如果它通常需要花费几跳将一个包从一个邻近 I_1 的位置传输到一个邻近 I_2 的位置，通过虫孔将 I_1 附近的包传输到 I_2 将会比通过网络中的多跳传输要提前到达。攻击者通过转发路由信息能使 S 和 D 相信它们是邻居，然后有选择性地丢弃数据包使它们间的通信中断。对于大部分路由协议来说，攻击者对于节点的影响也超出了虫孔端点的邻居。节点 S 将会宣告一个到达 D 的一跳路径，因此 C 将会通过 S 将包发送到 D。在几乎所有的基于需求的路由协议中，虫孔攻击能够通过隧道将路由请求信息直接发送到邻近目的节点的节点。因为路由请求信息通过一个高质量的隧道来传输，它比其他请求更早到达。根据协议，收到的为相同的路由发送的其他路由请求信息将会被丢弃。因此这个攻击防止了其他路由被发现，虫孔将会完全控制路由。攻击者能够建立一个 DoS 攻击来丢弃所有的信息，或者更加巧妙的，选择性地丢弃特定的信息来改变网络的功能[24]。一个有合适的虫孔的攻击者能够轻易地建立一个陷阱，攻击到达许多目的地的包（但不是转发）。一个智能攻击者可能能够有选择性地转发信息来激活其他的攻击，也能够将虫孔端点放到一个特定的位置。战略性地放置虫孔端点能够中断几乎所有的到达或者来自于一个特定节点和所有网络中其他节点的通信[24]。

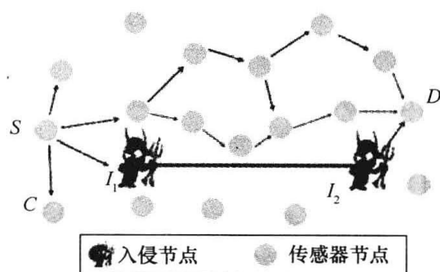


图 15-3　虫孔攻击：攻击者控制节点 I_1 和 I_2，并且通过一个低延迟链路将它们连接起来

周期性的邻居发现机制的路由协议严重地依赖于接收广播包作为邻居发现的一种方式，对于这种攻击也是相当脆弱的。

15.3.5 洪泛攻击

现存的基于请求的路由协议在每次路由发现中转发一个最先到达的请求包。在洪泛攻击中，攻击者利用路由发现操作这个特性，如图15-4所示。发起者节点发起一个到达目的节点的路由发现。如果被攻击者转发的这次路由发现的路由请求是第一个到达每个目标的邻居（如图中传感器节点），那么通过这次路由发现而发现的任何路由都将包括通过攻击者的一跳[25]。也就是说，当目标节点的邻居突然收到来自攻击者的请求时，它转发这一请求，但是不会进一步转发来自这个路由发现的请求。当非攻击的请求晚一点到达这些节点时，它们将会丢弃那些合法请求。因此，发起者将不能够发现任何包含至少两跳的（三个节点）可用路由（也就是说，那些不包括攻击者的路由）。

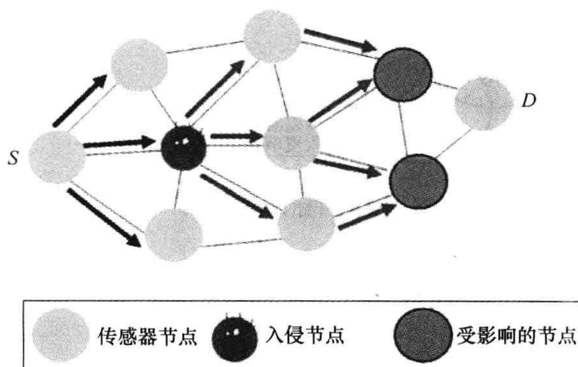


图15-4 网络说明洪泛攻击

一般地，一个攻击者能够比合法的节点更快地转发路由请求，因此它能够进入一个路由。这样的路由不能够很容易地被检测到。

15.3.6 伪装攻击

在这种类型的攻击中，一个系统实体非法伪装为另外一个实体来访问机密系统，也就是说，一个系统伪装成另外一个身份。假设一个节点 A 发送一个参考信号到它的两个邻居 B 和 C。一个攻击者 E 能够假装成 B，之后和 C 之间交换错误的时间信息，中断真正的 B 和 C 之间的时间同步处理，如图15-5

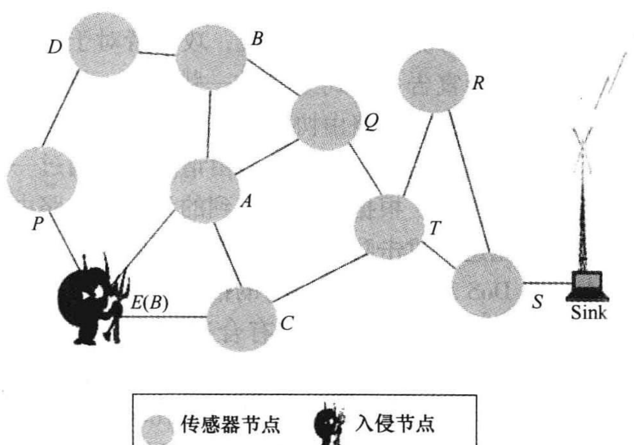


图15-5 网络说明伪装攻击

所示。

15.3.7 重放攻击

重放攻击是一种网络攻击，在这种攻击中一个有效数据传输将会被恶意地或者不正当地重复或者延迟。假设 Alice 想要向 Bob 证明她的身份。Bob 要求她的口令作为身份的证明，Alice 提供了这个口令（可能经过一些转换，像哈希函数）；同时，Eve 偷听了这次谈话，记住了这个口令。这次交互结束以后，Eve 伪装成 Alice 连接到 Bob；当要求一个身份证明时，Eve 发送上次偷听到的 Alice 的口令，Bob 一定会接受。

15.3.8 信息操纵攻击

对路由协议最直接的攻击是瞄准节点间交换的路由信息。通过欺诈、篡改或者重放路由信息，攻击者能够建立路由回路，扩展或者缩短源路由，如图 15-6 所示。这样的攻击者会攻击或者排斥网络通信，产生假的错误信息，划分网络，增加端到端延迟等。在这种攻击中，一个攻击者可能丢弃，修改，或者甚至伪造交换的时间信息来中断时间同步处理[26]。

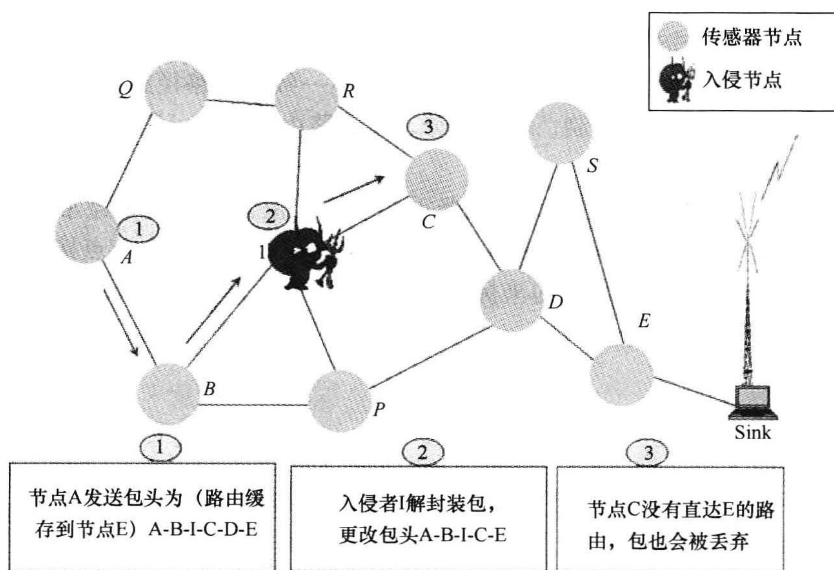


图 15-6 网络说明消息操纵攻击

15.3.9 延迟攻击

攻击者故意地延迟一些时间信息，例如在 RBS 模式中的信标信息，这样来使

时间同步处理失效。在图 15-7 中，入侵节点 E 在时间 t_1 收到信标信息，经过很长一段时间 t_n 后才发送它。

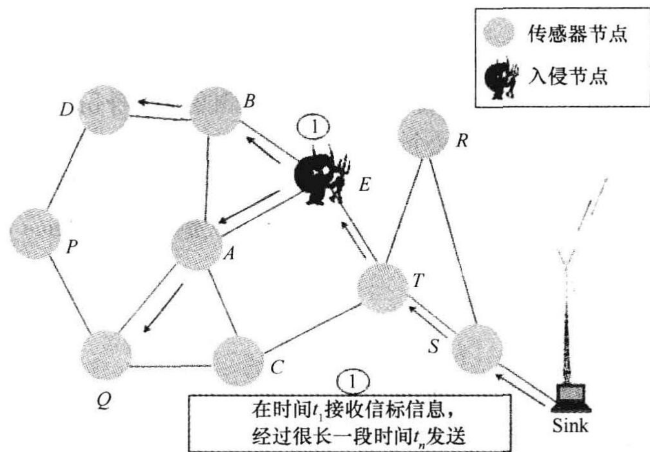


图 15-7 网络说明延迟攻击

15.3.10 Sybil 攻击

在一个 Sybil 攻击中，网络中一个单一节点扮演了多重身份。Sybil 攻击能够明显地降低容错模式的效力，例如分布式存储，不一致，多路径路由，和拓扑维护。被认为是用于分散节点的复制，存储划分，或者路由实际上可以使用一个单一的对手显示多重身份。

Sybil 攻击也对地理路由协议构成严重的威胁。位置感知路由经常需要节点与邻居之间交换协调信息来有效地路由地理地址包。仅仅期望一个节点接受，而不是一个来自它的邻居的单一系列的坐标，是合理的，但是利用 Sybil 攻击一个对手能够“一次出现在超过一个地方”。

15.4 反抗手段

这里，我们描述在传感器网络中可用的能够阻止攻击的反抗手段技术。将各种技术相结合能够阻止攻击，并且所有的技术都能满足一部分安全需求。在许多阻止攻击的方式中，仅仅一种技术不能够阻止所有的攻击；需要将一些技术结合起来。一些技术能够阻止被动攻击，一些能够阻止主动攻击。在这一章的所有技术中，我们考虑密钥管理、匿名通信和入侵检测。

15.4.1 密钥建立和管理

在传感器网络中，传感器节点是资源受限的，我们之前已经讨论过。这种资源

受限使得传感器网络的安全应用成为一个具有挑战性的问题。除了轻量级的密码外，也需要有效的密钥分发和管理机制。已经设计了许多密钥建立技术来应对有限的内存与安全之间的均衡，但是哪种模式才是最有效的仍然存在争议。这里，我们描述了一些基本的当前提出的协议。

15.4.1.1　单一广阔网络密钥、对偶密钥建立、受信任基站和认证

在参考文献[1]中描述了单一广阔网络密钥、对偶密钥建立和受信任基站。介绍了一种认证模式；尽管它不是一个密钥管理模式，但是它被许多模式所用。

15.4.1.1.1　单一广阔网络密钥

使用单一广阔网络密钥是到目前为止最简单的密钥建立技术。在这个技术的初始化阶段，一个单一的密钥预装入网络中的所有节点。经过部署，网络中的每个节点能够使用这个密钥来加密和解密信息。这项技术提供的一些优势包括最小存储需求和避免复杂的协议。节点的内存中仅仅存储一个单一的密钥，并且一旦部署在网络中，节点不需要执行密钥分发或密钥交换，因为所有的通信节点能够利用它们已经共享的密钥来进行信息转换。

尽管一个单一广阔网络密钥可能看起来是有利的，它的主要缺点是通过共享密钥单一节点的妥协将导致整个网络妥协。这种模式需要更少的计算和内存使用，但是也面对几个限制，因为它使得一个想要攻击的敌人对它进行攻击很容易，因此不能够提供一个传感器网络的基本需求。

15.4.1.1.2　对偶密钥建立模式

对偶密钥建立模式是传感器网络中最有效的密钥建立模式之一，因为与其他模式相比它不提供许多附加的特征，包括点到点认证和节点复制弹性。

对于一个采用对偶模式的有 n 个节点的网络，密钥的初步分配是通过为每个节点与网络中其他所有节点分配一个单独对偶密钥，也就是说，$n-1$ 对密钥，这些密钥存储在每个节点的内存中，因此每个节点能够同所有在其通信范围内的节点通信。由于每个节点同网络中的其他节点共享一个单一的密钥，这种模式提供了一个点到点的认证。每个节点能够证实与它通信的节点的身份。这种模式也提供了网络被俘获的适应力，因为一个被占领的节点不会泄露网络中不与他直接通信的节点的信息。通过增加适应力，这种模式最小化了节点被复制的概率。对偶机制的缺点是为每个节点同所有网络中的其他节点建立 $n-1$ 个单一密钥，以及在它的内存中维护那些密钥所需要的额外的花费。利用这样一个模式使得一个网络的大小是受限的，因为随着网络中节点数量的增加，每个节点内存中存储的密钥数量增加。如果一个网络有 10000 个节点，那么每个节点的内存中必须存储 9999 个密钥。因为传感器节点是资源受限的，这笔巨大的费用限制了这一模式的适用性，但是对于较小的网络这种模式是十分有效的。

15.4.1.1.3　受信任基站

利用对偶密钥建立模式的主要问题是网络中的每个节点需要存储 $n-1$ 个密钥

对。当我们使用一个受信任的基站来发送任意两个通信节点间的密钥时可以从根本上解决这个问题。这种模式也叫做集中式密钥分发中心（Centralized Key Distribution Center，KDC）方法。这种模式有很小的内存需求和一个完美的受控节点复制。它对于节点受控具有很强的适应能力，并且使得它可能取消密钥对。缺点是它是不可扩展的，基站成为攻击者的目标。

15.4.1.1.4　认证：μTESLA

加利福尼亚伯克利大学的Perrig等人[27]提出了一套针对传感器网络优化的安全协议，他们称之为“SPINS”。这套协议建立在两个安全建筑障碍上，每个执行独立的请求工作：SNEP和μTESLA。SNEP提供数据保密、认证、完整性和时新性，而μTESLA提供广播数据认证。μTESLA协议，用于规则网络中，被修改为SPINS用于资源受限的传感器网络中。SPINS在每个节点中包含TinyOS（操作系统），所有节点与基站通信。大部分WSN通信经过基站，并且包含三种通信类型：节点到基站、基站到节点和基站到所有节点。

SPINS协议的主要目标是设计一个基于SNEP和μTESLA的密钥建立技术来防止一个攻击者通过一个妥协节点向网络中的其他节点散播信息。这种模式中的每个节点同基站共享一个密钥，这个密钥在部署前进行了初始化。下面是说明这个模式如何工作的一些描述。

1）节点A和节点B是网络中的两个通信节点；

2）N_a由节点A产生；

3）X_{ab}是节点A、B间共享的控制密钥；

4）K_{ab}和K_{ba}是节点A、B间共享的加密密钥，它是由控制密钥X_{ab}产生的；

5）$\{M\}K_{ab}$表示利用密钥K_{ab}对消息M进行了加密；

6）$MAC(K'_{ab},M)$表示利用MAC密钥K'_{ab}为信息M计算MAC。

（1）SNEP：数据保密/认证/时新性：两种模式相结合形成SNEP，包括一个语义安全计数器和一个自举模式。使用这种结合，SNEP能够提供大量的优势，通过为每个信息仅仅添加8bit来降低网络的通信代价。它使用一个计数器，像许多其他协议一样，提供认证和时新性，但是这样做也使用提供语义安全的方法。我们应该注意到语义安全不是新的，它是一个密码学中通用的技术，例如利用传统的计数器（CTR）模式。节点间共享两个计数器，试图进行彼此通信，一些源节点的加密技术发送带有信息的共享计数器到目的节点。通常的加密可以作为一种简单形式的加密，但是对于保护信息是不足够的，然而，语义安全使得攻击者即使获得一个或更多的加密信息也难以取得原始数据，从而提供了更大的安全保护。在无线传感器网络中，用一个计数器发送信息会有开销，但是，能够通过节点间共享计数器并且每当目的节点收到信息时增加计数器来节省能量。和其他模式一样，为了更加安全，不应该重复使用同样的密钥。在SNEP中，使用单独的密钥来加密和进行MAC操作。源节点和目的节点的共享密钥用于为每个方向产生加密和MAC密钥。加密数

据有这样的形式 $E = D(K,C)$ ，D 是数据，K 是加密密钥，C 是计数器。MAC 是 $M = \mathrm{MAC}(K',C \| E)$。

在 SNEP 中，节点 A 发送到节点 B 的完整信息是

$$A \rightarrow B: D(K_{ab}, C_a), \mathrm{MAC}(K'_{ab}, C_a \| D(K_{ab}, C_a)$$

因为每次信息都是加密的，计数器的值也增加到一个不同的值，满足了语义安全性；因此，尽管相同的信息被加密，一个攻击者不能够解密信息。这与传统密码学中的 CTR 模式是完全一样的。

利用 SNEP，一个攻击者没有机会通过持续地发送计数器同步请求来执行 DoS 攻击，但是这可以通过对每个加密的信息发送计数器值或者绑定一个短的 MAC 到不依赖于计数器的信息上的方式来避免。数据认证可以通过 MAC 来实现。信息中计数器的值防止了一个攻击者重放旧的消息，重放旧的信息可能导致传感器网络的混乱和花费。由于计数器的值保存在通信的两个终端，ID 不随每个信息迁移，通信的花费可以忽略不计。计数器模式也允许弱的时新性。如果计数器的值被正确地核实了，它会泄露信息的顺序，但是仅仅保证了信息的顺序，并不能保证来自节点 B 的回复是由于节点 A 的信息引起的。为了获得强的包含延迟评估的时新性，信息中必须包含一个随机数。为了获得强的时新性，节点 A 伴随一个回复信息发送一个随机数 N_a 到节点 B ，节点 B 会用一个回复信息来重新发送随机数。这个过程能够通过含蓄地利用 MAC 通信中的现时来优化；因此，具有强时新性的整个 SNEP 协议是

$$A \rightarrow B: N_a, R_a \text{ and } B \rightarrow A: \{R_b\}_{(K_{ba}, C_b)}, \mathrm{MAC}(K'_{ba}, N_a \| C_b \| \{R_b\}_{(K_{ba}, C_b)})$$

如果 MAC 正确地证实，节点 A 将会知道来自节点 B 的回复是对它的信息的回复。在这种方法中，假设通信双方都知道计数器的值，因此不需要伴随每次信息发送这个值；尽管实际上，信息可能丢失或者受到干扰，导致计数器的值不一致。协议需要同步计数器的值包括用下面的方式自举计数器的值：

$A \rightarrow B: C_a$ with $B \rightarrow A: C_b, \mathrm{MAC}(K'_{ba}, C_a \| C_b)$ and $A \rightarrow B: \mathrm{MAC}(K'_{ab}, C_a \| C_b)$。

计数器的值不需要加密，因为协议需要强的时新性，两个通信方使用计数器作为随机数。MAC 不需要包含名字 A 和 B ，因为它们使用的密钥 K_{ab} 说明了哪个节点参与了通信。如果节点 A 意识到节点 B 的计数器 Cb 不是同步的，它可能为了达到强时新性随一个包含 N_a 的信息请求节点 B 的计数器，或者

$$A \rightarrow B: N_a \text{ and } B \rightarrow A: C_b, \mathrm{MAC}(K'_{ba}, N_a \| C_b)$$

（2）μTESLA：认证广播：认证广播数据是无线传感器网络中的一个关键问题，但是先前的针对这一问题的解决方法受到太多通信和计算代价的限制，因此，在资源受限的无线传感器网络中不是很有用。TESLA，作为这些策略之一，通过使用数字签名技术提供了一个带有认证的广播数据的无效模式，对每个信息添加 24 个字节的代价，而通常一个包仅仅分配 30B。因此，利用 TESLA 会导致几乎所有

的包空间被密码占用。并且，TESLA 随着每个它发送和接收的信息包泄露了密钥，这会大量地消耗无线传感器网络的能量。最后，TESLA 使用一个单向密钥链认证密钥，这些密钥不可能存储在每个节点中。Perrig 等人，通过一种包含不明显的花费的方式（即 μTESLA），修改了用于认证广播数据的 TESLA；这种方法降低了使用非对称机制的认证数据的能量需求。并且，不像 TESLA 每次发送或者接收一个数据包时会泄露密钥，μTESLA 在一个时间段内仅泄露一次。μTESLA 仅有的限制是它限制了认证发送者的数量，因为在一个传感器节点中存储单向密钥链是昂贵的。

μTESLA 能够通过一个延迟的同步密钥揭发来提供非对称加密类型的认证广播。为了广播基站与无线传感器网络中节点之间的认证信息，μTESLA 要求基站和节点是时间松散同步的，因此每个节点知道一个最大同步误差的上限。当基站想要向一个给定网络中的所有节点发送一个包时，它预先计算这个包的 MAC。因为网络的所有节点确定仅有一个基站能够计算 MAC，MAC 密钥在这个点上未及时地被揭露，因此它们不容易受到来自敌方的攻击。发送到节点的包存储在它们的缓存中，直到基站揭露相应的密钥。一旦密钥被揭露，密钥能够被使用单向函数 F 的节点认证。如果一个密钥是正确的，一个节点能够使用它来认证存储在它缓存中的包。

每个 MAC 密钥是由函数 F 产生的一串密钥序列。发送者随机选择链的最后一个密钥 K_n ，然后重复使用 F 产生单向密钥链。假设基站在时间段 t_1 已经发送了包 P_1 和 P_2 ，时间段 t_2 发送了包 P_3 和 P_4 ，时间段 t_3 发送了 P_5 ，时间段 t_4 发送了 P_6 ，接收到包的节点不能够马上证实它们的身份，因此节点将它们存储在缓存中。在特定的时间段发送的包使用与那个时间段相应的密钥认证。在这种情况下，使时间段之差为 2，接收节点与基站之间是松散的时间同步并且知道密钥 K_0 。假设所有发送关于包 $P_1 \sim P_5$ 的密钥信息的包丢失了，仅剩下装载关于包 P_6 的密钥信息的包到达了，接收节点仍能够通过由 P_6 提供的信息取得密钥来验证其他包的密钥。因此，尽管一些包可能丢失，节点仍能够利用接收到的密钥验证它们。为了达到这一目标，μTESLA 采用多个阶段，每个阶段执行一个专门的任务，包括发送建立，广播认证包，自举新接收者，以及认证广播包。

（3）关于 SPINS 的考虑：SPINS 使用更少的传感器节点内存，也就是说，当常规加密占用 20% 的空间，μTESLA 占用 574B，可接受的全部可使用的内存是 2KB[27]。这种模式的执行是高效的，因为无线传感器网络的带宽足够进行 SPINS 使用的原始加密。再者，大部分 SPINS 的设计可能在其他低端设备的网络中使用。最后，SPINS 的通信代价是很小的，具有安全的特征，例如数据时新性、认证和加密，仅在一个 30B 的包中增加了 6B 的代价，允许添加到每个包中[27]。当带宽和内存的限制稍微放宽时，SPINS 甚至能够提供更大的优势。

对于单个节点广播和认证数据不是那么容易，因为在一个节点的内存中存储一

个单向密钥是不可能的，使用一个函数计算密钥产生很多网络代价，每个节点同网络中其他的每个节点不共享一个公共的密钥。然而，这个问题有两种解决方法。第一种方法是，节点用基站传输所有需要向其他节点广播的数据。第二种方法，当基站使用单向密钥链产生认证密钥，节点广播数据到基站。因为无线传感器网络是资源受限的，在一个单一的阻塞密码中执行原始加密是高效的，因此，不能够负担额外的安全代价。然而，SPINS 需要一个强壮的加密原则。

1）阻塞密码：在无线传感器网络中使用 RC5 将会是高效的，因为它的规模小，并且高效。再者，作为一个算法，在许多攻击下它容易被监视。使用 TEA 也可以为阻塞密码工作，但是他不易受密码分析监视。DES 和其他算法由于传感器网络无法满足的大规模和高计算需求，对于阻塞密码是不可用的。

2）加密函数：阻塞密码的计数器（CTR）模式对于加密解密可以使用同样的函数，密码文档的大小在这种模式中是一样大的。当工作于 SPINS 的加密函数中，这两个特性使得这种模式在 SPINS 的加密函数中工作时非常有用。而且，CTR 模式也提供了语义安全，这是一个先前已经讨论过的强壮的加密性质。为了使用 CTR 模式，发送节点和接收节点必须在内存中维护一个计数器，如果需要的话，持有一个有效的方式进行计数器同步。在两端维护一个计数器的优势之一是信息将不需要一个自身装载计数器的代价。

3）时新性：使用一个计数器，并且每次发送一个信息它会自动增加，提供了弱的时新性。为了获得更强壮的时新性，发送者必须建立一个现时，并且应该将其包含到接收者的请求信息中。SPINS 使用一个 MAC 函数来产生随机数，并且建立了一个计数器来保存那些建立的路径。

4）消息认证：对于数据不仅需要一个好的加密函数，也需要一个安全的 MAC。因为阻塞加密会使用超过一次，CBC-MAC 用于 MAC。必须使用一个有效的消息建立方式来实现认证和消息完整性。结构$\{M\}_k, \mathrm{MAC}(k', \{M\}_k)$，$M$ 是数据，K 是加密密钥，K'是 MAC 密钥，它是安全的，避免节点破译不正确的密码文档。

这种模式的优势包括，它是已讨论过的最好的内存有效的模式之一，低复杂度，提供了强的安全特性，有一个全局设计允许将它应用于许多低端设备，低的通信代价，利用最小的代价提供认证和强的数据时新性。这种模式的不足包括经过一段特定的延迟，可能是信息延迟以及释放密钥的 μTESLA 花费。

15.4.1.2　公钥模式

加利福尼亚大学伯克利分校制造的 MICA2 节点有一个 8bit，7.3MHz 处理器，4KB RAM 和 128KB 的可编程 ROM[28]。大部分无线传感器网络使用同步密钥和其他非公钥加密模式[29]。这些模式的一个缺点是它们不像公钥密码体制一样灵活，但是它们计算速度更快。由于有限的内存、计算和通信能力，能量供应，传感器节点不能够采用精密的加密技术，例如典型的公钥加密。现有的实验还不足以将在无线传感器网络上使用公钥加密完全排除在外。通过使用 MICA2 节点和 TinyOS，来

验证公钥模式的执行性能。ECC比RSA[3]有更快的计算时间，更小的密钥，并且使用更少的内存和带宽。ECC和RSA都能够通过专用的协作处理来加速。最近，ECC被用于无线传感器网络[3,29]。作者在参考文献[3]中已经部署了一种方式来在无线传感器网络上执行公钥模式，这种方式是通过8MHz处理器，使用椭圆曲线计算加密密钥。MICA2节点在它8bit的处理器上也能够使用ECC。在这些传感器节点上，公钥能够使用1KB的RAM和34KB的ROM在34s内计算出来。下面的章节概述了RSA和ECC模式。

RSA和ECC模式

RSA和ECC模式都已经被研究了很多年了。它产生于1977年，仍然是当前可用的最流行的公钥加密技术之一。RSA依赖于分解非常大的数字的复杂性优势。ECC是由Koblitz和Miller两人于1985年独立发明的，它接近于公钥加密，基于椭圆曲线数学。ECC能够像RSA一样使用一个更小的密钥来获得同样的安全级别。一个160bit的ECC密钥和一个1024bit的RSA密钥有相同的安全级别[3]。一个224bit的ECC密钥与2048bit的RSA密钥相当[3]。这是由于它使用指数级的算法来解决椭圆曲线DL问题，这与在RSA中采用小的运行时间算法来解决大量的因数分解相反[3]。

RSA模式通常引入的密钥大小为512～2048bit。使用密钥K，它装载一个消息M，形成一个加密文档C。一个称为中国剩余定理的方法能够用于RSA加速。两个质数，q和p，相乘得到模数n[3]。计算这些模数倍增，CRT几乎能够降低四分之三的计算时间[3]。其他因素像蒙哥马利乘法和优化平方能够使RSA复杂性降低25%。

ECC由基于质数整数域或者比特多项式域在椭圆曲线上的点乘来计算的[3]。在无线传感器网络上执行ECC主要是对质数整数域感兴趣，因为比特多项式域数学不能够得到慢处理器的支持。ECC的操作线性扩展。这使得ECC比RSA在处理器上使用小字长上有优势。随着密钥规模增加ECC的优势也会增加。

在这些加密模式中，CRT、模块化乘法ECC和大的整数数学RSA是最重要的操作。由于处理器的小字长，大量的乘法操作需要高的内存读写。因此计算时间可以通过优化内在操作数量来降低[3]。

ECC在两个8bit的平台上执行。由于有限的资源，采用执行优化。部署了RSA-1024和RSA-2048来做比较[3]。结果表明具有一个私钥的ECC-160比RSA-1024更快。当将ECC-224与RSA-2048比较时，它的执行效果更好[3]。ECC，在两个平台上，比RSA-1024私钥操作要好。ECC由于处理器字长降低，与RSA相比，也改善了它的表现。

15.4.1.3 路由驱动椭圆曲线基于加密的密钥管理模式

在参考文献[6]中，基于ECC的密钥管理模式被描述为用于异构传感器网络。这里，作者划分了密钥管理模式为两个方面：集中式密钥管理和分布式密钥

管理。

1. 集中式密钥建立

在集中式的基于 ECC 的密钥管理模式中，一个服务器用于产生 ECC 公钥私钥对，一对密钥用于一个 L-传感器节点和 H-传感器节点。在按等级划分的分簇传感器网络中，L-传感器节点是叶子传感器，H-传感器节点是簇头传感器节点。服务器在一个大的有限区域 F 中选择一个椭圆曲线 E，并且在曲线上选择一个点 P。每个 L-传感器（标为 u）预先加载了它们的私钥（用 $K_u^R=I_u$ 来表示）。一个 H-传感器节点有一个大的存储空间，预先加载了所有 L-传感器节点的公钥（例如 $K_u^U=I_uP$）。每个 H-传感器节点（用 H 表示）也存储了每个 L-传感器节点和它们的私钥的关联。另外一个方法是预先加载每个 L-传感器节点的公钥，然后部署好后让每个 L-传感器节点发送它们的公钥到 H。然而，这种模式引入了大的通信代价，并且带来了安全问题，因为一个敌方可能在公钥路由到 H 的过程中对其进行修改。

在 H-传感器节点预先加载的密钥是由难以破坏的硬件来保护的。即使一个敌人占领了 H-传感器节点，它也不能够获得密钥信息。由于来自难以破坏的硬件的保护，同样的 ECC 公钥/私钥对能够用于所有的 H-传感器节点，这降低了密钥管理模式的存储代价。每个 H-传感器节点预先装载了一对公用的 ECC 公钥（用 $K_H^R=I_uP$ 表示）和一个私钥（用 $K_H^R=I_H$ 表示）。H-传感器节点的公钥 K_H^U，也在每个 L-传感器节点中加载了，并且这个密钥是用来认证来自 H-传感器节点的广播。ECDSA 算法[5]用于认证来自 H-传感器节点的广播。当 H 广播路由结构信息（例如 MST）到 L-传感器节点，利用 H 的私钥在这个信息上计算一个数字签名。每个 L-传感器节点能够使用 H 的公钥证实数字签名，因此认证广播。再者，每个 H-传感器节点预先加载了一个特殊密钥 KH，被一个对称加密算法使用，用于证实新部署传感器网络和保证 H-传感器节点之间的安全通信。

选择了簇头 H 后，每个 L-传感器节点 u，发送给 H 一个清楚的（未加密的）密钥请求信息，它包括 L-传感器节点的 ID—u，以及 u 的位置。可能使用一个贪心地理路由协议[40]转发密钥请求信息到 H。注意到，在簇形成阶段，簇头节点的位置对于所有的 L-传感器节点是已知的。一个 L-传感器节点发送信息到与簇头具有最短距离的邻居 L-传感器节点，下一个节点执行相似的操作，直到数据包到达簇头。

经过一段时间，簇头 H 会接收来自所有（或者大部分）簇内的 L-传感器节点的簇头请求信息，然后 H 使用一个集中的 MST（或者 SPT）算法来决定簇中的三个结构。然后，H 为每个 L-传感器节点和它的 c-邻居产生共享密钥。对于一个 L-传感器节点 u 和它的 c-邻居 v，H 产生一个新的密钥 $K_{u,v}$。记住 H 预先加载了所有 L-传感器节点的公钥。H 利用 u 的公钥和一个 ECC 加密模式[30]对 $K_{u,v}$ 加密，然后 H 将信息单播到 u。L-传感器节点 u 解密信息，获得它们之间的共享密钥。当所有 L-传感器节点获得共享密钥后，它们能够安全地同它们的 c-邻居通信。

2. 分布式密钥建立

密钥建立也可以采用分布式方式。在分布式密钥建立中，每个L-传感器节点预先加载了一对ECC密钥——一个私钥和一个公钥。当L-传感器节点（用u表示）发送它的位置信息到簇头H，u使用u的私钥计算一个信息上的MAC，这个MAC附加到信息中。当H接收到信息，H能够核实MAC，然后使用u的公钥验证u的身份。然后，H使用H的私钥为u的公钥产生一个证书（用CAu表示）。确定了一个簇的路由树结构以后，簇头H，传播树结构（也就是父子关系）和认证到每个L-传感器节点的相应的公钥。公钥证书标记了H的私钥，并且由于每个L-传感器节点预先加载了H的公钥，能够被每个L-传感器节点证实。一个公钥证书证明了一个公钥的真实性，进一步向另一个L-传感器节点证明了一个L-传感器节点的身份。

如果在路由树中的两个L-传感器节点之间是父子关系，那么它们是彼此的c-邻居，并且它们将自己建立一个共享的密钥。对于每对c-邻居，具有小的节点ID的传感器节点发起密钥建立过程。例如，假设L-传感器节点u和v是c-邻居，并且u与v相比有一个更小的ID。过程如下：

1）节点u发送它的公钥$K_u^U = I_u P$到v

2）节点v发送它的公钥$K_v^U = I_v P$到u

3）节点u通过倍乘它的私钥产生共享密钥I_u，利用v的公钥K_v^U，也就是说，$K_{u,v} = K_u^R K_v^u = I_u I_v P$；相似的，$v$产生共享密钥$K_{u,v} = K_v^R K_u^u = I_u I_v P$。

经过上面的处理，节点u和v共享一个公共密钥，它们能够开始安全通信。为了降低计算代价，对称加密算法用于L-传感器节点之中。注意到，在分布式密钥建立模式中，不需要假设在H-传感器节点中有难以破坏的硬件。

15.4.1.3.3 执行评估

这种模式的密钥存储空间能够像下面这样评估。假设在一个异构传感器网络（HSN）中，H-传感器节点的数量和L-传感器节点的数量分别是M和N。代表性地，假设$M \ll N$。在集中式ECC密钥管理模式中，每个L-传感器节点是预先加载了它的私钥和H-传感器节点的公钥。每个H-传感器节点预先加载了所有L-传感器节点的公钥，加上一对它们自身的私钥/公钥对和新部署节点的一个密钥K_H。因此，一个H-传感器节点预先加载了$N+3$个密钥。全部预先加载的密钥的数量是

$$M(N+3)+2N=(M+2)N+3M$$

根据网络中节点部署的阐述，部署前或者部署时，所有的L-传感器节点需要知道它们的H-传感器节点（L-传感器节点定义为一个低端传感器，H-传感器节点定义为一个高端传感器）和网络的树结构。因此，节点需要根据预先定义的树结构来部署，否则在L-传感器节点和与它们相对应的H-传感器节点之间会有一个不匹配的私钥/公钥对。这是部署过程中的一个笨重的程序。实际上，为了降低网络

中的 L-传感器节点的密钥空间，这种模式需要预定义网络的树结构。在那种情况下，假设网络中有 M 个 H-传感器节点和 N 个 L-传感器节点，每个 H-传感器节点预先加载了所有 L-传感器节点的公钥，加上它们自身的一对私钥/公钥对，以及另外一个对于新部署节点的密钥。每个 L-传感器节点预先加载了它的私钥和它的 H-传感器节点的公钥（L-传感器节点知道与它相应的 H-传感器节点，因为节点是根据一个预先定义的树结构部署的）。因此，一个 H-传感器节点预先加载了 $N+3$ 个密钥，一个 L-传感器节点预先加载了两个密钥。最终，全部预先加载了密钥数量是

$$M(N+3)+(2N)=(M+2)N+(3M)$$

这已经讨论过了。

如果这种模式不关心预定义的树结构那么每个 L-传感器节点需要预先加载 $M+1$ 个密钥。因此，全部预先加载的密钥数量是

$$M(N+3)+(M+1)N=2MN+3M+N$$

否则，这种模式会引入大的通信代价以及安全问题。这种模式是完全静态的因为预先加载了其他节点的密钥。每个节点需要维护一个同其他节点的相关链接。

在分布式 ECC 密钥管理模式中，每个 L-传感器节点预先加载了它的公钥/私钥。每个 H-传感器节点预先加载了一个公钥/私钥和密钥 K_H。因此，预先加载的密钥的全部数量是

$$3M+2N$$

在参考文献[54]中，Eschenauer 和 Gligor 提出了一个传感器网络中的基于概率密钥预分发的密钥管理模式。在这个模式中如果每个传感器节点预先加载了 m 个密钥。一个有 $M+N$ 个传感器节目的网络中所有预先加载的密钥数量是

$$m(M+N)$$

这种需要是完全静态的，因为预先加载了其他节点的密钥。每个节点需要同其他节点维护一个相关的连接。

15.4.1.4 基于身份和配对的安全的密钥管理模式

在参考文献［31］中，提出了基于身份和配对的密钥管理模式；这个模式将在下一节概述。

1. 协议设计

基站（sink 节点或者系统管理员）在网络自举期间有下面的额外的任务。

确定两个有同样主序 q 的群 G_1 和 G_2。我们把 G_1 看作是一个附加的群，G_2 看作是一个乘法群，如预备章节中所述。

确定双线性映射 $g:G_1\times G_1\rightarrow G_2$，防碰撞加密哈希函数 H_1 和 H_2，满足 H_1: $\{0,1\}^*\rightarrow G_1$，一个从随意长度字符串到 G_1 中的点的映射和 H_2: $\{0,1\}^*\rightarrow\{0,1\}^\mu$，

一个从随意长度字符串到一个 μbit 固定长度的输出的映射。

产生系统秘密

$$\omega' \in Z_q^*，这里 Z_q^* = \{y \mid 1 \leqslant y \leqslant q-1\}。$$

除了基站（sink节点）网络中的任何节点不知道 ω'。基站也使用这个秘密来产生非对抗节点的秘密点。

因此，系统参数 $\langle G_1, G_2, g, H_1, H_2 \rangle$ 被认为是非对抗节点。基站也为节点提供了下面的参数，关于它们的ID和秘密点。

单独提供每个节点（L-节点和H-节点）一个不同的，唯一的和真实的ID，让我们说

$$ID_{Ri} = H_1(ID_{R1}^{\ell}),\ ID_{R2} = H_1(ID_{R2}^{\ell}), \cdots, ID_{RN} = H_1(ID_{RN}^{\ell}) \in G_1$$

对于 N 个节点和相应的秘密点，$SP_{R1}, SP_{R2}, \cdots, SP_{RN} \in G_1$，定义为

$$SP_{Ri} = \omega' ID_{Ri} = \omega' H_1(ID_{Ri}^{\ell}),\ i = 1, 2, \cdots, N$$

也为每个节点提供一个不同的随机值 $R_{Ni} \in Z_q^*$，如果一个L-节点（叶子节点）不是直接连接到H-节点（头节点），那么它使用它的秘密节点和它的通信邻居的ID来产生它自身和它的通信邻居间的秘密共享密钥，也用来认证彼此（密钥产生和认证技术在下一节描述）。如果一个L-节点直接连接到H-节点，那么它使用它的秘密点来产生它和它的H-节点之间的秘密共享密钥，也用来认证彼此（密钥产生和认证技术在下节描述）。对于一个给定的元组 $\langle ID_R, SP_R \rangle$，没有一个能够决定系统的秘密 ω'，正如预备章节所讨论的。

为每个H-节点提供所有节点的 ID_s 和一个相应的随机数 $R_N \in Z_q^*$。这个随机数用于认证L-节点和与它相应的H-节点。基站同时改变这个随机数并将它发送到H-节点，这将在下一章讲述。

有了上面的信息，任务节点能够产生它自己的秘密共享密钥。让我们检查一个节点 K；K 从基站收到了它的 ID、ID_K 和相应的秘密点，$SP_M = \omega' ID_M = H_1(ID_M^{\ell})$。因此，$K$ 能够产生和它的通信邻居间的秘密共享密钥；让我们假设它的通信邻居是 M，它也收到了它的 ID、ID_M 以及相应的秘密点，$SP_M = \omega' ID_M = H_1(ID_M^{\ell})$。结果，$K$ 和 M 都能够产生它们自己的秘密共享密钥，而不需要共享它们的秘密点，如下，K 计算 $K_{KM} = g(SP_K, ID_M) = g(ID_K, ID_M)^{\omega'}$ 和 M 计算 $K_{MK} = g(SP_M, ID_K) = g(ID_M, ID_K)^{\omega'}$。根据2.1节的性质，这两个值 K_{KM} 和 K_{MK} 是一样的；因此不需要共享秘密，K 和 M 能够产生它们一样的秘密共享密钥。这些等式也具有先前在15.2.3和15.2.3节中引用的性质。因此，对于一个给定系列的 ID_s 和相应的秘密点 $\langle ID_K, SP_K \rangle$，没有一个能够决定系统的秘密 ω'。结果，一个节点能够产生它自己和它的通信邻居的秘密密钥而不需要共享它们相应的秘密点。

2. 密钥管理和认证

基站作为一个系统管理部门，负责为网络中的所有节点产生 ID_s 和相应的秘密

点。在网络自举时，任何节点知道自己的 ID 、秘密点和随机数。再者，H-节点知道所有 L-节点的 ID_s 和它们的随机数。节点的随机数在一个时间段内由基站改变，通过相应的 H-节点分发到 L-节点。因此，在开始阶段，每个节点有一个随机数，并且经过一段时间会改变。这个随机数用于 L-节点和它的通信邻居（邻居 L-节点）之间或者一个 L-节点和它的 H-节点之间的认证。

密钥建立的认证处理在下面的章节中介绍。

共享密钥建立和认证

经过网络自举和簇形成后，所有的 L-节点知道它们相应的直接连接的或者通过它们的通信邻居（其他 L-节点）连接的 H-节点。记住，所有的 H-节点彼此之间直接连接或者通过其他的 H-节点连接。此外，H-节点要么是直连连接到基站，要么是通过其他的 H-节点连接到基站。因此，节点是以分层的方式连接。

让我们考虑一个 H-节点和它的 ID、ID_{RH1}、相应的秘密点 SP_{RH1} 和随机的 R_{H1}；一个 L-节点和它的 ID、ID_{RL1}、相应的秘密点 SP_{RL1} 和随机的 R_{L1}。让我们再一次考虑，H-节点想要认证它的邻居 L-节点 ID_{RL1}。因此，H-节点产生它的秘密共享密钥 $K_{H1L1}=g(SP_{RH1},ID_{RL1})=g(ID_{RH1},ID_{RL1})^{\omega'}$ 和一个认证代码 $Aut0=H_2(K_{H1L1}\parallel ID_{RH1}\parallel ID_{RL1}\parallel 0)$ 发送到 ID_{RL1}。另一方面，L-节点产生它的秘密共享密钥 $K_{L1H1}=g(SP_{RL1},ID_{RH1})=g(ID_{RL1},ID_{RH1})\omega'$ 和一个确认代码 $Ver0=H_2(K_{L1H1}\parallel ID_{RL1}\parallel ID_{RH1}\parallel 0)$，将它与 $Aut0$ 比较；如果 $Ver0$ 与 $Aut0$ 相对称，那么它产生了另外一个认证代码，$Aut1=H_2(K_{L1H1}\parallel ID_{RL1}\parallel ID_{RH1}\parallel 1)$，并且将它发送到 H-节点。最后，H-节点计算 $Ver1=H_2(K_{H1L1}\parallel ID_{RH1}\parallel ID_{RL1}\parallel R_{L1}\parallel 1)$，并且将它与 $Aut1$ 比较；如果相配那么认证是成功的，否则是失败的。因此，所有的节点（H-节点和 L-节点）与它们的通信邻居或 H-节点证明了彼此。因此，网络中的节点能够彼此安全地通信。

当基站改变了 L-节点的随机数，它用基站和 H-节点间的秘密共享密钥加密这个随机数。因为 L-节点和它们相应的 H-节点是已经认证的，H-节点用 H-节点和相应的 L-节点间的秘密共享密钥来加密这个随机数，并且将它发送到 L-节点。如果 L-节点不是直接连接，而是通过其他 L-节点连接到他的 H-节点，那么 H-节点执行双重加密。在第一次加密中，它使用和目的 L-节点间的秘密共享密钥。在第二次加密中，它使用与它邻近的 L-节点间的秘密共享密钥，并且发送加密包到邻近的 L-节点。结果，L-节点解密这个包，查看目的 ID，然后再一次使用它的通信邻居（L-节点）的秘密共享密钥来加密这个包，再将它发送到 L-节点。因此，这个随机数最终到达了目的地 L-节点。我们注意到，中间 L-节点不能够获得关于随机数的任何信息，因为它被 H-节点双重加密了。

3. 安全分析和执行评估

本节介绍密码执行和密码分析协议。首先，从执行上看，然后介绍安全分析和密钥空间。

（1）密钥空间节约

在先前的协议中，节点是静态的，并且它们一点也不能够移动。因此，节点固定在一个确定的位置。而在这个协议中，节点能够移动或者它们能够被固定在一个位置；在这两种情况下，被提及的协议都是可行的。因为在现存的协议中节点是固定的，给出了当节点固定在一个特定的位置时密钥空间之间的比较。

为了与先前的协议进行比较，考虑两种场景。第一种场景是已经在参考文献[6]中提到了，也在先前的章节中讨论过了；在这种场景下，预加载密钥的全部数量是 $M(N+3)+2N=(M+2)N+3M$。第二种场景和第一种场景是一样的，全部的预加载密钥数量是 $M(N+3)+(M+1)N=2MN+3M+N$。

另一方面，对于这个协议，为了降低计算代价，没有预加载的密钥，簇形成后，密钥空间可能是 N/M；考虑节点在现存的模式中不移动。可以采用如下证明：在网络自举时，节点仅存储它们自己的秘密点；然后节点计算它们的邻居节点的秘密共享密钥。为了平等地划分网络中各个簇中的节点，每个簇中有 N/M 个节点，M 是H-节点的数量。为了降低计算代价，簇形成后，节点仅仅存储它们邻居的共享密钥；因此，更坏的情况下，需要存储的全部共享密钥数是 N/M。

（2）预防攻击

下面这些著名的攻击在传感器网络中是有效的，这个协议有效地预防了这些攻击。

1）伪装攻击：在这种模式中，所有的节点在它们的路由过程中彼此经过广阔网络认证。因此，入侵节点不能够伪装成一个有效节点，也不能够在有效节点间交换错误信息。所以，伪装攻击对于我们提出的协议是无效的。

2）重放攻击：在这种模式中，首先，一个攻击者节点不能够通过认证处理。此外，旧的确认或者认证代码经过一个特定时间段不再有效。因为基站改变了网络中的节点的随机数，而这个随机数是用来产生确认代码的，节点在通过过程中也执行加密。

3）信息窜改攻击：为了执行这种攻击，一个攻击者需要参与信息通信。最后，需要成为网络中的一个有效节点。在这种模式下，一个攻击者不能够冒充路径或者包。因此，这种攻击在这个协议中不是有效的。

4）延迟攻击：在这种协议中，节点能够计算从源到它们自身的距离，也能够估计包的传输时间。因此，如果任何一个入侵节点想要试图延迟一个包，那么它是不能够生效的。

15.4.2 匿名通信

本节介绍一些无线传感器网络中的匿名协议。匿名协议对于预防一些面向目标的攻击是有效的，尤其是主动攻击。

15.4.2.1 分层的匿名通信协议

在参考文献[32]中提出了分层的匿名通信协议（Hierarchical Anonymous Communication Protocol，HACP）。HACP 提供了两个不同的机制来达到匿名——一个是基于引入假消息来达到簇内匿名，另一个是基于一个基于环的方法来达到簇头节点间的匿名通信。

15.4.2.1.1 网络模型

这里，已经考虑了分簇的传感器网络，因为分簇允许 MAC 和路由的可扩展性。簇头也作为聚合点进行数据聚合，因此实际传输到基站的数据的数量降低了。将传感器节点分组，因此传感器节点仅与簇头通信，然后簇头将聚合信息发送到处理中心，可以节省能量。不同的环境下已经提出了许多分簇算法[33-36]。这些算法的目标是产生最小的簇使得任何簇中的任何节点离簇头至多 d 跳。

通信图 G(VCH,E)是用于表示分簇网络。VCH 是一系列的簇头，E 是一系列连接到簇头的通信边（可能是包含中间非簇头节点的路径）。我们假设 G 是被连接的。

图中一个生成树最初是固定的。然后，使用图中生成树的一个欧拉遍历（DFS 遍历）定义一个环。环信息能够使用潜在的路由协议来获得能量有效性和负载均衡。

这个协议是基于对称密钥加密技术，因为在传感器网络中部署公钥协议是不可行的[27]。有许多传感器网络的密钥预分配模式来建立节点间的密钥[37]。在这个协议中，假设每个传感器节点和它的簇头之间共享一个密钥。每个簇头和环中的邻居簇头共享一个对称密钥。下面的标记用于描述这个协议：$E(M,K_{ij})$ 表示用 K_{ij} 加密信息 M，$D(M,K_{ij})$ 表示用 K_{ij} 解密信息 M，K_{ij} 是结点 i、j 间的共享密钥。

令牌和帧。任何时候仅有一个帧穿越这个环。节点使用一个令牌传递证访问机制来访问期间一个经过网络的帧。一个想要发送数据的节点应该首先接收许可。当它得到了通行证的控制权时，它可以在那个帧中传输数据。每个帧有固定的长度，并且包含了令牌自身的状态。一个令牌要么是空闲状态，要么是被占用状态。帧的格式如下

$$\langle E((Token \,\|\, E(Message_{Header},K_{sd}) \,\|\, E(Message_{Data},K_{sd})),K_{si})\rangle \tag{15-1}$$

式中，K_{si} 是源节点 s 和节点 i 间的共享的秘密密钥，节点 i 是发送者 s 的上行邻居；K_{sd} 是源节点 s 和目标节点 d 共享的秘密密钥。

令牌的形式如下：

$$\langle Redundancy\ predicate \,\|\, Status\rangle$$

Redundancy predicate 是用于检查帧的有效性。为了使帧能够成功地被节点 i 确认，必须实现解密 *Redundancy predicate*。“Status” 指定了令牌是 “被占用” 还是 “空闲的”。如果令牌是空闲的，一个节点能够通过那个帧发送数据，否则不能够

发送数据。

$Frame_{Header}$ 的形式如下：

$$\langle Redundancy\ predicate \parallel Source\ Address \parallel Destination\ Address \rangle$$

FrameData 的形式如下：

$$\langle Data\ length \parallel Data\ Padding \rangle$$

Data length 指明了包中全部数据的长度。当需要发送的数据的数量不足以填充整个帧时这是十分重要的。在那种情况下，发送的数据被填充了一些随机数来满足固定长度的帧的大小限制。

1. 簇中匿名通信

在一个网络中插入假的通信是一种隐藏网络内部的通信模式的技术，这使得通信分析更加困难[38]。假的通信的产生增加了通过混合网络发送的信息的匿名性。

一个假信息是指由一个传感器节点建立的假信息。最终目的地是它的簇头；假信息被簇头丢弃。网络观察者和其他节点不能够从真实信息中区分出假信息。

在 HACP 中，每个传感器节点（包括簇头）以一个泊松频率 r_t 传输信息。因此，平均来说，每个传感器节点每 $1/r_t$ 时间内发送一个信息。r_s 表示每个传感器节点的感知频率。因此，无论何时有感知数据需要发送，传感器利用它与簇头之间共享的密钥来解密数据信息并发送它。否则传感器节点发送假信息。因此，假信息以一个频率（$r_t - r_s$）发送。当一个簇头有一个信息想要发送到一个它的簇节点时，簇头简单地利用它与传感器节点共享的秘密密钥对信息进行加密然后发送它。当一个传感器节点感知到一个包的传输时，它接收这个包，用它的密钥解密，然后检查它是否是一个有效的包。

2. 簇头节点间的匿名通信

无论何时一个节点 i 接收到一个帧，它使用它与环中的下行节点共享的密钥解密这个帧，确认冗余谓词。一旦满足了 Redundancy predicate，执行下面的算法。

1）如果节点没有数据要发送，它仅仅利用与它的上行节点间共享的普通密钥来解密作为结果的简单帧，然后将包重传到环上。

2）如果令牌的状态是空闲的，节点有一些数据需要发送到另外一个节点 D，然后 i 建立帧如下：

① 节点 i 建立 $Frame_{Header}$ 和 $Frame_{Data}$，正如先前解释的，使用与目的地之间共享的密钥。

② 节点 i 将令牌中的状态域设为被占用状态。

③ 使用与上行节点间共享的密钥计算公式（15-1），并且将这个包传输到环上。

3）如果令牌的状态被设置为被占用，节点使用共享密钥解密 $E(Frame_{Header}, K_{sd})$，检查帧中的数据是否是以自身为目的地的，以及检查是否满足 Redundancy

predicate。

① 如果节点能够检查帧头部的有效性，那么它会被传送到节点 i，节点 i 会复制它。然后利用与上行节点之间共享的密钥加密整个帧，并将这个帧传输到环上。

② 另外，如果节点 i 不能够检查帧头的有效性，那么不会传送给它，节点仅利用与上行节点之间共享的密钥来加密整个帧，并将这个帧传输到环上。

一旦帧返回到源节点，只要它有数据要发送，源节点重复这个过程。当它不再有数据需要发送时，它将令牌的状态域置为空闲，并将整个帧分配给一些随机产生的数据。然后它使用与上行节点之间共享的密钥来加密整个帧，并将帧传输到环上。

15.4.2.1.2　多重环

在一个由 n 个节点组成的网络中，环的大小是 n。这样，一个消息需要沿着整个环传输，因此，每个消息传输 n 次。为了降低通信代价（复杂度），协议将图划分为子图，在每个子图中建立环。在参考文献[27]中描述了同样的分割机制。一旦子图的划分是可用的，那么每个子图中的每个环也是可用的，这是由子图的生成树的欧拉遍历形成的。作为多个环上的一部分的节点被叫做交叉点。每个子图中至多有 x 个节点；因此，每个子图中的时间复杂度不超过 x。

为了激活与子图外的一个节点之间的通信，每个环分配了一个唯一的身份——RID。每个节点也知道目的地所属的环的 RID。我们在帧中引入了一个新的头节点——$E(Frame_{RID},K_{sJ})$ 来识别目的地的 RID，K_{sJ} 是源节点与交叉节点之间共享的普通节点，也是一个环的一部分，需要通过它到达目的地。这个帧修改后的形式如下

$$\langle E((Token \parallel E(Frame_{RID},K_{sJ}) \parallel E(Frame_{Header},K_{sd}) \parallel E(Message_{Data},K_{sd})),K_{si})\rangle$$

$Frame_{RID}$ 的形式是 $\langle RIP \parallel RID_D\rangle$。$RIP$ 是需要被满足的冗余谓词，标示成功地解密。RID_D 是目的地环的环标识。发送者利用与交叉节点共享的密钥来解密 $Frame_{RID}$，交叉节点是环到达目的地环的一部分。

当一个环中的一个节点有数据要发送到另一个环中的一个节点时，帧需要从一个环传输到另一个环，直到它到达了目的地的环。为了这样，每个交叉节点维护一个转发路由表，这个表指定了到达一个特定的目的环的帧需要被转移到哪个环上。上面的一个交叉节点成功地解密 $E(Frame_{RID},K_{sJ})$，存储帧的一个复本，然后重传这个帧。基于目的地节点的 RID 的交叉节点决定帧将会被转移到哪个环上。然后，它等待其他需要传输帧复本的环上的一个空闲令牌，使用它与到达目的环的下一个交叉节点共享的普通密钥对这个帧加密。这个过程持续到帧到达目的地环，接收到帧的 RID_D 的交叉节点仅仅分配一些随机的字符串给 $E(Frame_{RID},K_{sJ})$，并将帧传输到环 RID_D 上。

这种机制防止了本地通信遍历整个网络。即使一个敌方能够盗用一个交叉节点，他仅能够知道这个帧想要到达的环，不能够知道其他信息。攻击者甚至不能够

搞清楚帧的起源环。因此，这个机制不能减少协议提供的匿名。在一些情况下，仅仅一些节点可能需要匿名，这样的话仅需要在那部分节点中建立一个环。在这种情况下，网络中的一个环中的邻居并不一定为物理上的邻居，并且这些节点能够使用可用的最短路径进行通信。

15.4.2.1.3　HACP的工作性能

本节中，HACP的工作性能用强加的代价和提供的匿名来表示。

1. HACP的通信代价

在HACP中，当一个节点有数据需要发送时，它捕获一个空闲令牌，在那个帧中发送数据。另外，它仅转发空闲的帧。通信代价用于表示与空闲帧相应的通信数量。应该注意的是一个传感器节点的能量消耗可以从当前平均消耗[39]来求出，表示为

$$I_{avg} = T_{on}I_{on} + (1 - T_{on})I_{stby}$$

式中，T_{on}为接收机或者发射机是打开的时间片；I_{on}为当接收机或者发射机是打开时电池的当前能耗；I_{stby}为当接收机和发射机都关闭时的当前的电池能耗。因此，提高通信代价便会提高T_{on}，这意味着更高的能量消耗。

考虑一个有N个节点的环，其中有N_a个节点想要以单位时间内R个包的频率发送数据。让我们假设一个帧能够在单位时间内遍历环的最大时间为t。t的值依赖于环延迟，反过来依赖于帧的传输时间（T_{tr}），环遍历时间延迟（T_t）和一个节点的处理延迟（T_{proc}）。这里，可以忽略一个节点在转发一个帧前对它的处理延迟。因此，$t = 1/(NT_{tr} + T_{proc} + T_t)$。

如果环中出现了n个令牌，那么环中最多可以传输nt个帧。因此，理想情况下，似乎需要满足下面的条件，才能够做到没有空闲的帧被传输：

$$(N_a/N)R = n^*,t$$

因此，通过环的空闲帧的百分比是$1 - \{(N_aR)/(Nnt)\}$。因此，通信代价，也就是，空闲帧相应的传输量，由空闲帧的数量×环中的节点数量表示$= N - \{(N_aR)/(nt)\}$。随着环中节点数的增加通信代价几乎呈线性增加。

2. 安全分析

传感器网络以数据为中心的行为使得它们容易受到通信分析，事件位置识别以及活动区域的攻击。因此，确保数据匿名是一个关键的研究领域。使用通信分析，攻击者能够将数据流模式与事件位置/活动区域关联起来盗用网络功能。上面讨论的协议是HACP，它隐藏了节点的位置，并且掩盖了事件区域和来自窥探敌方的数据流之间的关联。隐藏来自窥探敌方的节点，协议阻止了信息-操纵攻击、延迟攻击、重放攻击，也阻止了节点遭受被动攻击。这个协议提供了匿名度和通信延迟代价之间的灵活的均衡。

15.4.2.2　在匿名传感器网络中寻找路由

在参考文献[40]中讨论了一个问题以及相应的匿名路由发现的算法。算法总

结和模拟结果在下面的章节给定。

15.4.2.2.1　问题和算法

假设有 n 个传感器节点随机部署到一个二维空间中，存在一个单一 sink 节点。传感器节点没有身份证明，甚至不知道它们有空间中的坐标。sink 节点广播一个单一的问题到所有的传感器节点，n 个传感器节点中的 n^* 个需要应答。我们称这 n^* 个传感器节点中的每个传感器节点为一个源，假设源是随机分布在这 n 个节点中的。这里需要应对的问题是寻找从所有源到 sink 节点的路由。因为源仅能够以一个低的能量广播，它们的应答可能不能直接到达 sink 节点，而是需要通过其他传感器节点路由。所有的 n 个传感器节点，即使它们中 $n-n^*$ 个不需要对 sink 节点应答，也可能需要参与聚集和中继源的应答信息。

这种方法潜在的关键前提是每个传感器节点能够测量在 sink 节点的传输中它进行感知所消耗的能量。对于传感器节点 i，它用 P_i 表示这个测量值。明显地，如果对于传感器节点 i 和 j，$P_i > P_j$，那么 i 比 j 更靠近 sink 节点，假设 sink 节点的广播各向均等地到达所有的传感器节点（也就是说，离 sink 节点同样的距离会消耗相等的能量）；尽管没有方法将传感器节点区分开来，通过使用这样一个性质，可能提供来自所有源到达 sink 节点的路由。

首先，引入一个简单的由 sink 节点和传感器节点执行的分布式算法。假设 sink 节点知道一个从 sink 节点到传感器节点最大距离的上界 R。如果 B_0 是 sink 节点广播它的询问的能量等级，那么传感器节点 i 当广播到达时，测量 P_i，能够计算它到 sink 节点的距离，距离用 R_i 表示，也能够测量它到以 sink 节点为中心的半径为 R 的圆形的径向距离（也就是 $R-R_i$）；它的所有代价是进行伴随着它的询问的 sink 节点广播 B_0 和 R 的值。让 T 和 T_i 分别表示一个相离为 R 和 $R-R_i$ 的电磁波的传播时间（这些能够简单地通过给定的在考虑中的波在中介中的传输速度来计算）。最初由 sink 节点广播，传感器节点需要做的除了接收这个消息以外，也包括 sink 节点或者一个传感器节点接收到来自一个传感器节点的消息时如何作出反应。算法的描述中，S_0 和 S_i 用于表示所使用的数据结构，相应地，由 sink 节点和传感器节点 i 来聚集它们收到的所有信息。如果传感器节点 i 是一个源，那么假设 S_i 开始包含它对 sink 节点的询问的应答。

下面的描述是关于行为 1 和 2 的，分别针对 sink 节点和一个普通传感器节点 i。行为 2，特别地，依赖于两个参数 f 和 r 的乘积 fr。f 和 r 分别是区间 $[0,1]$ 之间的一个数和一个传感器节点的广播半径是指节点能够直达的区域。一旦知道了 r 的值，假设传感器节点以一个和所有传感器节点一样的能量等级进行广播，因此，接收消息的位置正是那些与传感器节点之间的距离不超过 r 的位置。

1）行为 1：sink 节点广播 $Question(B_0,R)$，设一个计时器在经过 $2T$ 时间后中断。同时，一旦收到一个消息 $Answer(*,S)$，sink 节点将 S 并入 S_0 中。当计时器过期时，源发送到 sink 节点的应答都汇合到了 S_0 中。

2）行为2：一旦接收到消息 $Question(B_0,R)$ ，如果传感器节点 i 是源就广播消息 $Answer(P_i,S_i)$ ，并且无论它是否是源都要设一个计时器经过 $2T_i$ 个时间单元后中断。同时，一旦接收到一个消息 $Answer(P,S)$ ，传感器节点 i 将 S 并入 S_i 中，并打上合适的标签 P 。当计时器过期，传感器节点 i 检查 S_i 是否有来自于一个 *Answer message* 的信息需要合并到它中。如果有，它从 S_i 中所有 P 标记满足 $P < P_i$ 的实体中选择 P 标记最大的。如果选择是成功的（也就是说，至少有一个候选实体），那么让 P_j 作为最大的 P 标记；然后，传感器节点 i 计算来自 P_j 的 R_j 。如果不成功，那么传感器节点 i 让 $R_j = \infty$ 。如果 $R_j - R_i > fr$ ，那么它广播 $Answer(P_i, S_i)$ 。

15.4.2.2.2 模拟结果

对于每次模拟，n 个传感器节点随机部署在圆内，然后随机选择 n^* 个源。假设一个传感器节点在一次模拟中的每次广播能够准确到达那些以发射节点为中心的半径为 r 的圆内。r 的值是确定的，因此期望的圆内的传感器节点密度和在一个半径为 R 的更大的圆的密度是一样的。如果 n_r 表示预计的半径为 r 的圆内的传感器节点数量，那么 $n_r/\pi R^2 = n/\pi R^2$ ，因此它满足

$$r = \sqrt{n_r/n} * R \tag{15-2}$$

参数 r 是 n_r 的一个函数，因此，在实验中这两个参数用 f 和 n_r 来区别。对于

$$n = 2000 \text{ , } n^*/n = a \times 10^{-b}$$

式中，$a = 1,2,5$ 并且 $b = 1,2,3$ ；$f = 0.1, 0.3, 0.5$；$n_r = 9,11,13,15$（根据式（15-2），n_r 的这些值对应于 $r \approx 0.067R, 0.074R, 0.081R, 0.087R$）的结果。

模拟结果显示，当远离发射 sink 节点时传感器节点感知的放射状的能量消耗能够支持随机部署的源到达 sink 节点的路由结构。

15.4.2.2.3 安全分析

在这个协议中，考虑了匿名传感器网络，描述了对于从一组传感器针对一个 sink 节点的查询传递信息建立路由的问题。为了应对由匿名所带来的种种限制，在来自 sink 节点的查询中，协议依赖于使用传感器节点感知的能量等级。因此，对外部节点掩盖了节点的 ID_s 阻止了被动攻击。

15.4.3 入侵检测

入侵检测系统（Intrusion Detection System，IDS）是计算机网络中的重要的安全工作。在传统计算机网络和无线自组织网络中有许多方法应对入侵检测问题[41]，但是关于传感器网络这一方面的文献是很少的。故障恢复的目标是通过重启或者重编程失效或者出故障的节点来延长传感器网络的寿命。在结合中，这两个测量数据提高了一个潜在攻击者的花费。即使一个攻击者试图捕获一个节点，为达到自己的目的滥用这个节点，这个节点的异常行为很可能会被检测到，发现了这个节点，这

个攻击便是无效的。

当试图防止一个系统遭受恶意使用时，重要的是定义目标和潜在的攻击者的能力。需要考虑到那些试图通过控制节点执行的代码来占用节点的攻击者。这允许一个攻击者参与网络的普通操作，因此会影响到网络的操作结果以及占用节点的资源。这里不考虑 DoS 攻击。

一个攻击者向一个节点中注入恶意代码有许多种方法，包括利用它的应用代码或者使用的应用管理协议的弱点，以及物理上的脆弱点。软件脆弱点的影响可以通过使用优良的保险工具，像代码确认和其他的工具来最小化。保护物理脆弱点试图重写应用代码需要设置障碍，使得访问节点的硬件物理特性尽可能地困难[42]。然而，所有的防御机制增加了传感器网络的成本。因此，将资源用于入侵检测和发现来减轻攻击者的影响是明智的。

主动测量来抵御物理控制也是可能的。已经成为感知节点一部分的传感器节点能够帮助检测物理控制。例如，一个节点被重新部署了，加速度传感器节点能够发起关键处理，使得该节点在网络中无效。原则上，可以绕过所有防御机制，但是需要的代价是相当高的。大体上，我们想要避免这种情况，如果许多节点已经被攻击了，那么攻击一个单一节点变得更加廉价。

这里，讨论了一个 emotional-ant-based 方法来识别可能的预攻击行为，其次它与一个集中式的入侵检测机制相一致。在传感器网络中安全监视是通过自然界蚂蚁群居觅食行为来实现的。蚂蚁可能位于相互连接的传感器网络中的一个相关的位置，在本节中对于一些相关的需要描述的相关词汇请见参考文献[43,44]。所提出的方法的一个重要的优势在于很容易得到入侵者-遍历路径。

15.4.3.1　使用情感蚂蚁的传感器网络上的入侵检测

从上面的讨论，我们知道有许多针对入侵检测的技术。在参考文献[45]中，已经提出了在传感器网络上使用情感蚂蚁（Intrusion Detection on Sensor Networks Using Emotional Ants，IDEA）进行入侵检测，这个协议如下。

15.4.3.1.1　蚁群方法

入侵检测的数据挖掘方法首先使用于入侵检测自动化模式的审计数据挖掘中。几个应用于审计数据的数据挖掘算法计算模型能够精确地捕获入侵的实际行为和正常活动[46]。审计数据分析与挖掘将联系原则与分类算法结合起来发现审计数据中的攻击。其他的方法包括知识的模糊规则库系统分类器[44]、遗传编程技术[43]、支持向量机和决策树[Ant7]。在参考文献[47]中分析了一个分等级的分布式 IDS 结构。在参考文献[48]中，引入了一个自组织的基于蚂蚁群体的分簇技术（ANTIDS）来检测入侵。在设计 IDS 中，已经设想了不同的自适应和自组织技术。这里，一些相关的工作专注于一个不同的形式，这里蚂蚁代理将会建立一个框架，允许使用者定义一个给定的由于入侵产生的相互影响的特征。蚂蚁系统，在参考文献[49,50]中给定，下面介绍对于 IDS 的情感蚂蚁方法。

15.4.3.1.2 蚁群算法系统

一个代理 x 可以用元组 $\langle\beta,S\rangle$ 来描述，β 是一个代理的信任，S 是一个代理的状态。一套代理 A 实际上组织为一个家族来方便整个代理家庭之间的合作与通信。不必说的是，所有代理共享一个目标来发现在一个优化方法中的入侵行为。蚂蚁代理通常借助于信息素的浓度差异包含于这种分布式系统中。信息素是指蚂蚁能够释放到环境中的化学物质。信息素通过环境传播（通过布朗情感）也会随着时间挥发。

信息素相对简单的机制证明了一个简单但是有效的分布式的决策支持机制。其基本算法如下：

```
Initialize pheromone values (τ)
  while termination condition not met do
    for j = 1 to k do
      S^j ← construct solution (τ)
    end for
  Apply online delayed pheromone update (τ, s^1, ..., s^k)
  end while
```

在 initialize pheromone values 阶段，基本上初始化所有的信息素的值为相同的正的常数值，满足下面的条件：

1）无论一个节点是否已经被蚂蚁访问过，维护了一个内存（叫做 tabu 表）。它在一个特定的遍历中被扩展，然后在访问之间置空。

2）距离的相反数 $\eta_{ij}=1/d_{ij}$ 叫做能见度。能见度严格地基于位置信息，表示了当一只蚂蚁在节点 i 上时倾向于选择节点 j 的可能性。能见度用于指导蚂蚁的搜索能力，尽管一个基于只使用它的建设性的方法能够产生一个非常低质量的解决方法。

3）虚拟信息素轨迹的量 $\tau_{ij}(t)$，处于边缘并且是线上更新的。

Apply online delayed pheromone update $(\tau,s^1,\ldots,s^k)$ 用于存储踪迹和 tabu 表中的边缘细节，使用下面的信息素更新原则：

$$\tau_j \leftarrow (1-\rho)\,\tau_j + \sum_{j=1}^{k} \Delta s^j \tau_j \tag{15-3}$$

式中，如果 s^j 促成 τ_j，$\Delta s^j\tau_j=f(s^j)$，否则它为0.

$\Delta s^j\tau_j$ 是一个解 s^j 与信息素更新值 τ_j 的结合；k 是用于更新信息素的解的数量；ρ 是蒸发率；f 是一个函数，通过将一个解的质量映射为它的相反数。

15.4.3.1.3 情感蚂蚁系统

基本的思想是通过调查信息素的浓度来识别传感器网络中受到入侵干扰的路径。这个工作也强调了代理的情感方面，它们通过信息素更新在它们自身间交流一个特定路径的特征。因此，在一个传感器网络中，如果蚂蚁（这里称为情感蚂蚁）

被部署好了，它们能够保存在网络路径中轨迹的改变，遵守一个描述了大概可能性攻击的特定的原则的知识基础。一旦节点间的特定路径被间谍情感蚂蚁检测到，它能够通过信息素配平与其他蚂蚁交流路径特征；然后，向网络管理员报告。整个模式由几个同一时期的工作激发。在单纯的认知形式中，这种方法最终合并了两个基本部分：

1）情感和它在决策支持中的功能；

2）蚂蚁代理转化为情感蚂蚁。

首先，讨论了一般代理的情感模型。任何情感模型的结构主要是基于由真实世界的代理显示的对某个特定主题的响应。它们围绕喜好、满意、沮丧和靠近[Ant20]展开。

1. 情感代理定义

这里，需要设想和建立代理（定义特征，例如名字和类别）。一个代理可能是一个对象类型，它们的功能是与环境相关的，或者非对象类型，它们的功能与代理必须执行的行为有关。

2. 原则基础

对于可能的攻击场景的原则，正如网络管理员的提议，是累计的。再者，可以提出某些原型来规划传感器网络的这些原则。

3. 情感模板

这些模板代表个人和不同的情感状态，也就是说，如果参数与现存的代理相符。对于一个清楚的情感交换模型的图片，在下面的章节里形成了一些数学概念。

4. 蚂蚁情感模型

这项工作精确地采用蚁群算法[49,50]。在蚁群算法中，只有那些从路径开始产生最好的遍历的蚂蚁才会被允许进行对枝上的信息素浓度进行全局更新。

15.4.3.1.4　安全分析

保护一个传感器网络的流行方式是采用加密技术或者通过保护来自未授权访问/操作的感知信息，以及实施有效的入侵检测机制。这里讨论的方法是一个基于蚁群的入侵检测机制，这个机制也能够跟踪入侵者。这项技术可以与传统的基于机制学习的入侵检测技术相结合来保护传感器网络。上面讨论的情感蚂蚁模型提出了合作的分布式的智能，作为一个面对不确定性，有软时限的不完整信息以及资源限制的分布式合作问题。IDEA 模型框架的一个重要特征是感知行为模式的能力，审议，以及由可能值发起的基于自组织原则的行为，以此来检测网络中的恶意节点。因此，它防止了所有基于信息注入的攻击行为，例如信息操纵攻击、重放攻击、延迟攻击等。

15.4.3.2　在无线传感器网络中应用入侵检测系统

一次入侵可以定义为能够导致一次未授权访问或者改变一个特定系统的一系列行为。IDS 的任务是监测计算机网络和系统，检测可能的网络入侵，以及检测到入

侵后报告给用户，如果可能的话重新配置网络。一个无线传感器网络由一系列的能够维护彼此的无线通信信道，而不会依赖于任何固定基础设施的节点组成。这个因素以及其他因素使得无线传感器网络本质上不同于有线网络。对于传感器网络，IDS也必须向基站发送警报来提醒人类用户。最后，IDS必须是简单的，对于指定的传感器网络威胁以及在网络中使用特定的协议能够做出高度专业化反应。参考文献[51]中的一个大体的IDS结构将在下面的章节中描述。

15.4.3.2.1 传感器网络的大体的IDS结构

1. 检测实体

传感器网络内在的限制，如稀少的资源和有限的电池寿命，要求对于如何执行监测任务进行周密的计划。与在一个自组织网络中，IDS代理必须位于每个节点中一样。然而，考虑到工作性能，这些代理的结构必须分为两部分：本地代理和全局代理。

1）本地代理应该监测本地活动，发送的信息以及传感器接收的信息。这种监测活动仅当传感器活动时执行，并且传感器仅仅管理它自己的通信。因此，传感器节点上的负载是低的。

2）全局代理应该观察它们的邻居的通信，也能够作为看门狗。然而，并不是所有节点都能够同时执行这个操作，因为这个操作需要传感器分析它们通信范围内的所有包的内容。因此，一次必须仅仅是一个特定的节点子集观察网络通信。

一旦一个代理，全局的或者本地，发现网络中的一个可能的安全破坏，它必须建立和发送一个警告给用户。用户只有通过基站才是可达的。因此，所有的警告必须发送到基站。这种发送警告到基站的机制依赖于基本的传感器网络的结构，但是它必须确定所有警告安全地到达了它们的目的地（使用的机制如μTesla）。

2. 数据结构

每个代理，也就是每个节点，为了恰当地工作必须存储关于它周围事物的信息。这些信息可以分为两类：关于安全的信息（一个警告数据库，包含了关于警告和可疑节点的信息）和关于环境的信息（一系列的与节点直连的邻居，它们能够使用接收到的信息，在节点的寿命时间以内更新）。

每个节点有一个内在的警告数据库，用于存储节点代理产生的安全信息。这个数据库的形式和规模是依赖于具体实现的（也就是说，它们依赖于在传感器网络中使用的协议）。然而，它们必须包含下面的域：建立时间、类别和警告源。

如果任何部署信息是可用的，或者部署以后，使用与LEAP[52]协议相同的假设（网络将在开始部署后的ts是安全的），这个邻居列表可以提前得到。这个列表的一个问题是它的内存足迹。这个列表中一跳邻居的数量像一个二次函数一样扩大，$((n^2)+n)$，因此，对于高密度网络它是不可扩展的。然而，列表的规模可以通过使用Bloom filters来降低[53]，用1bit的邻居列表数组存储每个邻居的邻居列表。对于一个$k=1$（哈希函数），并且$m/n=2$（比特数两倍于邻居数）的配置，列表

的规模降低了75%，引入一个16% ~ 40% 的错误定位代价。

3. 本地代理

本地代理的任务是通过分析信息的本地资源发现任何影响传感器节点正常行为的攻击或者威胁。这些资源是节点的实际状态，也就是节点接收到的包和发送的包，对环境的测量数据和所有可用的邻居节点的信息。

哪种攻击应该被本地代理检测到？首先，如果节点能够知道它们自身是否被操控了，对于传感器节点物理或者逻辑安全的攻击是可以被检测到的。例如检测它们是否被重编程了，因此，传感器节点能够在允许任何新代理执行前发起一个警告。

节点测量数据也是易受攻击的。因为传感器节点的主要任务是分析环境数据，任何敌方能够试图为它自己的利益影响这个过程。然而，所有的被节点读取的数据来自真实世界，遵守特定的模式和限制。因此，可以用异常检测技术来监视这些测量数据。例如如果一个节点将要部署到一个静态区域，加速表中的任何变动意味着节点被一个未知源吸引，因此它会发出一个警告。最后，本地代理也监测直接到节点的包。然而，关于本地处理的包有一些问题是独立于协议的，例如一个新节点加入网络中以及信号干扰。

在静态场景中，初始部署后很少有节点会加入到网络中，每个节点能够利用已知的邻居列表。每次一个节点接收到来自一个新邻居的包，它将把它加入列表中，发起一个警告。如果网络上的人类用户知道他们没有向网络加入任何节点，他们将得知新节点属于敌方。如果一个节点试了很多次发送一个包，但是信道不可用，可以采用误用技术来检测这是否是一个需要发送警告的非正常情况。

4. 全局代理

正如先前所描述的，全局代理必须负责分析它的直连邻居发送或者接收的包。它们也能够作为一个看门狗，使用独立于协议的技术接收和处理由下一跳节点中继的包。因为全局代理能够接收来自邻居的由下一跳节点中继的包（由于通信的广播性质），它们能够通过分析这些包来监测一个特定的节点是否在丢弃或者修改包。

然而，所有的全局代理是活动的，并且同时监听它们的邻居，分析网络必将是一个对于能量来说代价高的操作。结果，只有覆盖传感器网络中所有通信的传感器节点的一个特定的子集应该发起它们的全局代理。这项任务如何完成是依赖于传感器网络的潜在结构的。有两个基本的结构指定了传感器节点如何路由网络中的信息以及传感器节点如何进行自身分组。这两个结构称为分层的和平面的。

在分层配置中，将传感器节点分簇。簇中的一个成员作为服务器，或者一个“簇头”（簇头不比簇中其他成员更加强大），具有管理和路由的任务。另一方面，在平面配置中，信息由一个传感器节点路由到另一个传感器节点（网络中的每个传感器节点参与路由协议），几乎所有传感器节点有同样的计算能力和限制。

在分层结构中，全局代理在每个簇头中发起，因为将所有的簇头连接起来

（在大部分情况下）会覆盖整个传感器网络。因此，整个网络覆盖得到了保证。这种方法帮助维护系统的整体能量，因为簇头要么是比其他节点更加强大，或者是节点之间周期轮流地扮演簇头。

这个分簇方法在建立阶段和最大全局覆盖的簇维护中增加了网络复杂度，增加了一个可能的攻击点和周期控制信息代价。然而，有另外一个可以用于平面结构的分布式方案，不需要将它们分簇或者增加一些更强大的节点，称之为自发看门狗。

5. 自发看门狗

自发看门狗技术依赖于传感器通信的广播性质，利用区域中部署的传感器节点的高密度性。对于网络中循环的每个包，有一系列节点能够接收这个包和下一跳的中继包。因此，所有的这些节点有机会发起它们的全局代理来监测那些包。主要目标是在网络中循环的每个包仅仅发起一个全局代理。过程如下：

1）由于通信的广播性质，每个活动节点将会接收它的邻居发送的所有包。

2）节点将会检查它自身是否是包的目的地。如果不是，它不会立即丢弃这个包，而是检查包的目的地是否在它的邻居中（这样它能够接收由目的地转发的任何包）。因为所有节点存储了一个所有邻居节点的邻居列表，因此可以执行这个检查。

3）如果节点自身为目的地，节点能够作为一个自发看门狗。因此，它将计算网络中有多少节点处于与它相同的情况。

4）如果满足需求的节点的数是 n，一个单一节点将会以 $1/n$ 的概率选择它自身作为这个包的全局代理。这个过程类似于 n 个人和一个有 m 个面的骰子，$n = m$，试图在骰子中获得一个1来激发全局代理。

15.4.3.2.2 安全分析

上面的章节讨论了一个通过优化观察传感器通信来检测异常和监视包的技术。这样，一个节点的任何可疑行为可以被检测到，并发送到基站；最后，进行阻止这种恶意行为的行动。这样，外部节点不能够参与重编程，也不能够伪装自身。因此，能够防止消息操纵攻击、sybil攻击、伪装攻击和虫孔攻击。

15.5 总结

在传感器网络中，传感器节点是资源受限的；因此，对于传感器网络来说安全应用是具有挑战性的问题。在这一章中，我们关注于关于传感器网络安全的攻击以及与它们相对应的应对措施技术。仅使用一个技术不能防止所有的攻击。因此，需要把它们结合起来。根据资源受限和应用来选择安全协议。

除此以外，为了理解协议，我们也讨论了数学背景和一些其他的相关准备工作。我们主要关注于最新提出的协议和一些基本协议。所有单独协议的安全分析在它们相应的子章节中都单独给出了。没有协议能够防止所有的攻击。而且到目前为止，还没有一种方案是针对传感器网络安全中的DoS攻击的。

参考文献

1. Y. Xiao, V.K. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway, A survey of key management schemes in wireless sensor networks, *Computer Communications Journal*, 30(11–12), 2314–2341, 2007, Elsevier, Science Direct.
2. N. Koblitz, Elliptic curve cryptosystems, *Mathematics of Computation* 48, 203–209, 1987.
3. N. Gura, A. Patel, A. Wander, H. Eberle, and S.C. Shantz, Comparing elliptic curve cryptography and RSA on 8-bit CPUs, in *Proceedings of the 6th International Workshop on Cryptographic Hardware and Embedded Systems*, Boston, MA, August 2004.
4. N. Koblitz, *A Course in Number Theory and Cryptography*, 2nd edn., *Graduate Texts in Mathematics*, vol. 114, Springer-Verlag, New York, 1994.
5. I. Blake, G. Seroussi, and N. Smart, *Elliptic Curves in Cryptography*, London Mathematical Society, Lecture Note Series 265, Cambridge University Press, Cambridge, U.K., 1999.
6. X. Du, M. Guizani, Y. Xiao, S. Ci, and H.H. Chen, A Routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks, *IEEE Transactions on Wireless Communications*, Accepted for publication, Apr. 2007 (to appear). Extended version of, A routing-driven key management scheme for heterogeneous sensor networks, *IEEE International Conference on Communications*, 2007 (ICC'07), ISBN: 1-4244-0353-7, 24–28 June 2007, Glasgow, DOI: 10.1109/ICC.2007.564, pp. 3407–3412.
7. A. Shamir, Identity-based cryptosystems and signature schemes, in *CRYPTO'84: On Advances in Cryptology*, pp. 47–53, Springer-Verlag, 1984.
8. R. Sakai, K. Ohgishi, and M. Kasahara, Cryptosystems based on pairing, in *Symposium on Cryptography and Information Security* (*SCIS2000*), Okinawa, pp. 26–28, January 2000.
9. L.B. Oliveira, D.F. Aranha, E. Morais, F. Daguano, J. Lopez, and D. Ricardo, TinyTate: Computing the tate pairing in resource-constrained sensor nodes, in *6th IEEE International Symposium on Network Computing and Applications* (*NCA 2007*), Cambridge, MA, pp. 318–323, July 12–14, 2007.
10. L.B. Oliveira, D. Ricardo, J. Lopez, F. Daguano, and Loureiro, Identity-based encryption for sensor networks, in *5th IEEE International Conference on Pervasive Computing and Communications Workshops* (*PerComW'07*), White Plains, NY, pp. 290–294, March 2007.
11. L.B. Oliveira, D. Aranha, E. Morais, F. Daguano, J. L'opez, and R. Dahab, TinyTate: Identity-Based Encryption for Sensor Networks, available at http://eprint.iacr.org/2007/020.pdf.
12. J.L. Hill and D.E. Culler, Mica: A wireless platform for deeply embedded networks, *IEEE Micro*, 22(6), 12–24, 2002.
13. P. Levis, S. Madden, J. Polastre, R. Szewczyk, K. Whitehouse, A. Woo, D. Gay, J. Hill, M. Welsh, E. Brewer, and D. Culler, TinyOS: An operating system for wireless sensor networks, in W. Weber, J. Rabaey, and E. Aarts, eds., *Ambient Intelligence*, Springer-Verlag, New York, 2004.
14. The Tate Pairing, available at http://www.computing.dcu.ie/ □ mike/tate.html.

15. Elliptic Curve Cryptography Tutorial, http://www.certicom.com/index.php/ecc
16. D. Boneh and M. Franklin, Identity based encryption from the weil pairing, *SIAM Computing*, 32(3), Extended Abstract in Crypto 2001, 586–615, 2003.
17. S.L.M. Berreto, H.Y. Kim, and M. Scott, Efficient algorithms for pairing-based cryptosystems, in *Advances in Cryptology-Crypto'2002, LNCS 2442*, Springer-Verlag, Berlin, Germany, pp. 354–368, 2002.
18. S. Galbraith, K. Harrison, and D. Soldera, Implementing the tate pairing, in *Algorithm Number Theory Symposium-ANTS V, LNCS 2369*, Springer-Verlag, Berlin, Germany, pp. 324–337, 2002.
19. D. Boneh and M. Franklin, Identity-based encryption from the weil pairing, in *Advances in Cryptology—CRYPTO'01, Lecture Notes in Comput Science*, vol. 2139, Springer-Verlag, Berlin, Germany, pp. 213–229, 2001.
20. D. Boneh, B. Lynn, and H. Shachum, Short signatures from the weil pairing, in *Advances in cryptology—ASIACRYPT'01, Lecture Notes in Comput Science*, vol. 2248, Springer-Verlag, Berlin, Germany, pp. 514–532, 2001.
21. A. Joux and K. Nguyen, Separating decision Diffie–Hellman from Diffie–Hellman in cryptographic groups, Cryptology ePrint Archive, Report 2001/03, available at http://eprint.iacr.org/2001/03/.
22. R. Dutta, R. Barua, and P. Sarkar, Pairing-based cryptographic protocols: A survey, Cryptology ePrint Archive, Report 2004/064, available at http://eprint.iacr.org/2004/064.
23. Y.C. Hu, A. Perrig, and D.B. Johnson, Packet leashes: A defense against wormhole attacks in wireless ad hoc networks, in *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies* (*INFOCOM 2003*), San Franciso, CA, 2003.
24. Y.C. Hu, D. Johnson, and A. Perrig, SEAD, secure efficient distance vector routing for mobile wireless ad hoc networks, in *IEEE Workshop on Mobile Computing Systems and Applications*, Callicoon, NY, June 2002.
25. H. Y-Chun, A. Perrig, and D.B. Johnson, Rushing attacks and defense in wireless ad hoc network routing protocols, in *WiSe '03: Proceedings of the 2003 ACM Workshop on Wireless Security*, San Diego, CA, ISBN: 1581137699, pp. 30–40, September 19, 2003.
26. X. Du, M. Guizani, Y. Xiao, and H.-H. Chen, Secure and efficient time synchronization in heterogeneous sensor networks, *IEEE Transactions on Vehicular Technology*, 57(4), 2387–2394, July, 2008.
27. A. Perrig, R. Szewczyk, V. Wen, D. Culler, and D. Tygar, SPINS: Security protocols for sensor networks, in *Proceedings of 7th Annual International Conference on Mobile Computing and Networks MOBICOM 2001*, Rome, Italy, July 2001.
28. D. Malan, M. Welsh, and M.D. Smith, A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography, in *Proceedings of 1st IEEE International Conference Communications and Networks* (*SECON*), Santa Clara, CA, October 2004.
29. A.S. Wander, N. Gura, H. Eberle et al., Energy analysis of public-key cryptography for wireless sensor networks, in *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications* (*PERCOM*), Kauai, HI, March 2005.

30. B. Karp and H. T. Kung, GPSR: Greedy perimeter stateless routing for wireless networks, in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, pp. 243–254, 2000.
31. Sk. Md. M. Rahman, N. Nasser, and K. Saleh, Identity and pairing-based secure key management scheme for heterogeneous sensor networks, in *Proceedings of 4th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2008)*, Avignon, France, October 2008.
32. A. Durresi, V. Paruchuri, M. Durresi, and L. Barolli, A hierarchical anonymous communication protocol for sensor networks, in *Proceedings of 2005 IFIP International Conference on Embedded and Ubiquitous Computing (EUC-05)*, Nagasaki, Japan, LNCS 3824, Springer-Verlag, pp. 1123–1132, December 2005.
33. S. Bandyopadhyay and E.J. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies* (*INFOCOM 2003*), San Francisco, CA, March–April, 2003.
34. C.F. Chiasserini, I. Chlamtac, P. Monti, and A. Nucci, Energy efficient design of wireless ad hoc networks, in *Proceedings of European Wireless*, February 2002.
35. O. Younis and S. Fahmy, Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach, in *Proceedings of IEEE INFOCOM'04*, Hong Kong, March 2004.
36. O. Younis and S. Fahmy, HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks, *IEEE Transactions on Mobile Computing*, 3(4), 366–379, October 2004.
37. W. Du, J. Deng, Y.S. Han, S. Chen, and P. Varshney, A key management scheme for wireless sensor networks using deployment knowledge, in *Proceedings of IEEE INFOCOM'04*, March 2004.
38. C. Daz and B. Preneel, Reasoning about the anonymity provided by pool mixes that generate dummy traffic, in *6th International Workshop, Information Hiding (IH'04)*, Revised selected papers, *Lecture Notes in Computer Science* (LNCS-3200), Springer-Verlag, Toronto, Canada, May 2004.
39. E.H. Callaway, Jr., *Wireless Sensor Networks: Architectures and Protocols*, Auerbach Publications (an imprint of CRC Press), New York, 2003.
40. R.C. Dutta and V.C. Barnosa, Finding routes in anonymous sensor networks, arXiV:cs.NI/0507021, vol. 1, Jul 7, 2005, EUC 2005, *LNCS 3824*, pp. 1123–1132, 2005.
41. Y. Zhang and W. Lee, Intrusion detection in wireless ad-hoc networks, in *MOBICOM 2000*. ACM Press, 2000.
42. R. Anderson and M. Kuhn, Low cost attacks on tamper resistant devices, in *5th International Workshop on Security Protocols*, Paris, France, LNCS 1361, Springer-Verlag, pp. 125–136, April 1997.
43. A. Abraham, Evolutionary computation in intelligent web management, evolutionary computing in data mining, in *Studies in Fuzziness and Soft Computing*, A. Ghosh and L. Jain, eds., Springer Verlag, Germany, Chapter 8, pp. 189–210, 2004.
44. A. Abraham, R. Jain, S. Sanyal, and S.Y. Han, SCIDS: A soft computing intrusion detection system, in *6th International Workshop on Distributed Computing* (*IWDC 2004*), A. Sen et al. eds. Springer Verlag, Germany, *Lecture Notes in Computer Science*, vol. 3326, pp. 252–257, 2004.

45. S. Banerjee, C. Grosan, A. Abraham, and P.K. Mahanti, Intrusion detection on sensor networks using emotional ants, *International Journal of Applied Science and Computations*, 2005.
46. D. Barbará, J. Couto, S. Jajodia, and N. Wu, ADAM: A testbed for exploring the use of data mining in intrusion detection, in *ACM SIGMOD Record*, 30(4), 15–24, December 2001.
47. S. Marsella, W.L. Johnson, and C. LaBore, Interactive pedagogical drama, in *Proceedings of 4th International Conference on Autonomous Agents* (*ICMAS*), 2000.
48. V. Ramos and A. Abraham, ANTIDS: Self-organized ant-based clustering model for intrusion detection system, in *4th IEEE International Workshop on Soft Computing as Transdisciplinary Science and Technology*, Muroran, Japan, May 2005.
49. J. Deng, R. Han, and S. Mishra, INSENS: Intrusion-tolerant routing in wireless sensor networks, Technical Report CU-CS-939-02, Department of Computer Science, University of Colorado, November 2002.
50. M. Dorigo and L.M. Gambardella, Ant colony system: A cooperative learning approach to the travelling salesman problem, *IEEE Transactions. Evolutionary Computation*, 1, 53–66, 1997.
51. R. Roman, Z. Jianying, and J. Lopez, Applying intrusion detection systems to wireless sensor networks, in *3rd IEEE Consumer Communications and Networking Conference* (*CCNC 2006*), vol. 1, pp. 640–644, January 2006.
52. S. Zhu, S. Setia, and S. Jajodia, LEAP: Efficient security mechanisms for large-scale distributed sensor networks, in *10th ACM Conference on Computer and Communications Security* (*CCS'03*), Washington, DC, October 2003.
53. B. Bloom, Space/time trade-offs in hash coding with allowable errors, *Communications of the ACM*, 13(7), 422–426, July 1970.
54. L. Eschenauer and V. D. Gligor, A key management scheme for distributed sensor networks, in *Proceedings of the 9th ACM Conference on Computer and Communication Security*, Washington, DC, November 2002.

第 16 章　无线传感器网络中的网络管理技术

近年来，由于无线传感器网络（WSN）扩展了跨越全球以及进入我们日常生活的各种网络设备的可用性和商业潜力，同时具有持续的可用性，引起了研究团体的大量关注。为了保持这种网络一直运转，需要健壮的网络管理。传感器网络管理的范围包括监测，或者控制一个网络，但是它缺少在应用和易损性方面的发展而导致内在和诱导的失效。本章中，我们研究了无线传感器网络中不同寻常的网络管理技术，特别是在结构，设计和应用方面。我们对经典的网络管理框架提出更加新颖的方面，并且对于出现的各种应用场景强化它们的优点。

16.1　概述

无线传感器网络（WSN）包含大量的智能传感器，通过提供一种新的环境监测和控制方式来转换先前从环境中收集的数据的范例。由此产生的一系列传感器网络应用包括医疗、军事、家庭自动化、环境监测以及工业监测与控制。

传感器网络在许多方面不同于传统网络，显示了一些独有的特征。传感器节点通常被制作成很小的尺寸（cm^3 或者 mm^3），这个尺寸限制导致严重的资源限制，例如有限的电池能量，低计算能力和内存资源，贫乏的无线带宽，以及有限的通信能力。然而，如果 WSN 的网络资源得到良好的管理，整个网络的工作性能和网络收益是很可观的。

一个网络的监测、管理和控制操作过程可以被看做是“网络管理”。更具体地说，这样的管理功能对网络中的所有部件协调配置、安全、操作和维护。然而，网络管理对于每个网络需要采用不同的方法。这样，大量的传统管理方法由于它们自身的特性，不能够直接应用于传感器网络中。无线传感器网络的应用设计成在它们的内在的能量、带宽和资源限制下工作，与传统网络相比，它们用于达到更好的工作性能和响应时间。事实上，错误通常发生在大规模的 WSN 中，有成百上千的节点和严重的资源限制。因此，在这种情况下，部件维护和给节点再充电是不可行的，传感器节点通常被认为是一次性的。再者，参考文献[3]的许多调查研究表明，环境干扰和配置错误甚至能够导致在整个 WSN 开始操作以前造成损耗。

再者，传感器节点经常部署于严酷的环境中，这增加了 WSN 节点构造的显著变化的可能性。因此，一个传感器网络管理系统必须允许一个网络在没有预先知道网络拓扑的情况下的自形成，自组织以及自配置。对于自管理的 WSN，当节点以一种粗略的方式部署的情况下，节点自举并执行一个自测试[12]。然后，它们发现

事件位置并观察能量级别，使用状态以及管理情况。所有这些活动由自管理功能在网络管理层执行。如果节点发现了它们的位置，它们能够将它们自身划分为组。尽管传感器网络管理的意义相当重大，WSN管理不存在通用的方法[1]。

再者，传统的或者临时网络用于大量的用户应用。这意味着加载和配置网络部件来支持各种类型的服务。另一方面，WSN总体上是面向应用的。与传统网络和临时网络相反，无线传感器网络协同运行一个通用应用软件。

为了应对WSN的这些行为特性，需要设计一种具有新类型的管理功能的网络管理系统。本章将会探讨和提出这样一个用于监测和控制WSN的管理框架。

16.2节回顾WSN管理的设计目标，紧接着16.3节讲述管理规模。16.4节描述管理结构的其他设计方案。16.5节讨论一些当今在WSN管理领域的研究成果。16.6节将*IP - USN*描述为一个整合技术，紧接着16.7节将WSN管理作为一个FCAPS（Fault Configuration Accounting Performance and Security）模型。16.8节对这章进行总结。

16.2 WSN管理的设计目标

当设计用于传统网络中时，管理应用可能在吞吐量和延迟上有限制，同样的，当设计用于无线传感器网络中时会存在硬件限制，如内存、能量和处理能力。传感器网络的网络管理目标需要与面向应用的传感器网络设计建立一个清楚的和直接的关联。目标定义见下面章节内容。

16.2.1 可扩展性

假设WSN中的节点是大量部署的，新的节点可能动态加入网络中。网络管理结构应该能够与网络中密集部署的节点以及产生的大量数据一起工作。

16.2.2 有限的能量消耗

传感器设备几乎全部靠电池运行。管理操作在一个节点的资源上应该是轻量级的来延长它的寿命；采用这种方式，有助于延长网络的整体寿命。管理操作应该是高效的，并且应该消耗尽可能少的能量。

16.2.3 内存和处理限制

传感器设备都是有一个受限的内存和有限的处理能力。管理应用对于管理信息应该强加更少的存储代价。应该优化信息基站和被管理的代理来迎合内存和处理限制；它们应该有一个有限的处理代价和代码足迹来最佳地使用网络资源。

16.2.4　有限的带宽消耗

当出现高的信道损伤时传感器网络可能会出现带宽限制，所设计的管理应用应该考虑到这点。在这种情况下，可能需要对庞大的管理查询进行优化来满足连接特性。再者，与通信相关的能量代价通常超过感知和处理代价。因此，考虑到带宽限制和大量与传输相关的能量消耗，管理查询应该被优化。

16.2.5　网络动态适应性

对于网络动态性和拓扑改变，管理应用应该是可调整的。它们也应该能够聚集网络当前的状态和拓扑变化。像增加了新的节点，当前节点失效，以及网络内部和外部的节点移动性，可以根据应用需求能够被这些网络所支持。

16.2.6　容错性

传感器网络中的错误不同于那些传统网络中的错误。一个错误可能是由一个处于睡眠状态以节省能量的节点引起的，也可能是由一个能量耗尽或者由于移动性或者网络划分而与网络断开的节点引起的。管理系统和应用应该表现出自愈特征，并且它们应该知道这样的网络动态。

16.2.7　网络应答

网络中的事件应该被快速报告给网络管理系统。网络动态，例如节点移动，一个节点的失效，网络划分子网络，必须被管理系统迅速地检测到。管理系统应该能够响应和适应这些变化，并根据管理策略进行调整。

16.2.8　设备代价

管理系统应该通过在网络中部署最少的用于管理应用的设备来最小化部署代价。管理系统应该重用当前的网络硬件和有效地利用它的资源用于管理操作。

16.3　管理规模

经典的网络管理范例定义为两层，也就是说，管理功能和管理层。管理功能覆盖了故障管理、结构管理、计费管理、性能管理以及安全管理。它们一起合称为 FCAPS 模型。管理层定义了网络管理范围的抽象。

16.3.1　管理功能

传统上，WSN 被认为是面向任务，以应用为中心的，但是一些最近的提议为它定义了新的角色，例如，在一个 WSN 上运行多个应用和以服务为中心的性质。

多重应用的执行，从服务来看，赋予 WSN 一个完全不同的应用前景。例如，部署于感兴趣区域的一个 WSN 能够执行各种各样的军事应用，包括目标跟踪，挖掘检测和识别一个朋友或者敌人。

所有这些情况要求在传感器节点上管理功能必须是自治地执行。因此，对于 WSN 管理一般需要一个自治网络管理范例。这里，我们简单地描述了基于自身的自治的网络管理。

16.3.1.1 自管理

自管理可以描述为“通过管理对象管理工作网络”。自管理功能将部件转化为复合的实体来最优地提供所需要的功能。它使得系统根据一个管理员的目标利用最少的人为干预来管理它们自身。自管理的实现能够通过构造，优化，自愈，计费和安全模型来更好地理解，如图 16-1 所示。我们也应该注意到自管理策略的需求使得特别是结构管理变得更加重要。故障、性能、计费和安全管理，都起源于结构管理并最终汇聚到结构管理。

16.3.1.2 自配置

一个部件的自配置决定了它的运行和维护特性以及应用执行，数据通信和数据转发。自配置模块使管理部件根据高层策略自动进行自配置。

16.3.1.3 自愈

自愈描述了每个节点有能力察觉到它运行时的不正常现象，在无人干预的情况下，做出必要的调整来恢复到正常运行状态的性质。

16.3.1.4 自计费

自计费能够对网络的使用和其他本地资源进行测量。

16.3.1.5 自安全

自安全应该能够最小化对所有部件以及网络资源的未授权或者意外的访问。安全管理功能处理保证合法使用，维护保密性等。

16.3.1.6 自优化

自优化意味着持续地寻找提高网络工作性能的方法。节点不断识别和抓住机会使得整个网络的工作性能更高效。

16.3.2 管理层

对于逻辑结构设计，管理功能依赖于它们所属的各自的管理层。定义管理层和框架的方法可以分为“top-down”和“bottom-up”方法。前者提高了清晰度和粒度，因为分析上层有助于定义下层的需求。相反地，后面的方法是易于操作的。管理层分为任务、服务、网络和网络元素层。

16.3.2.1 任务层

任务层表示任务对象，任务对象主要处理服务进展和代价受益分析。这些目标决定一个 WSN 基本骨架的主要资源，并描述了其蓝图的基本特点。这一层特别关

注于网络建立，维护，感知，处理以及通信代价，概述了传感器网络的轮廓。

16.3.2.2　服务

一个 WSN 服务是与应用目标相关的功能。在设计这些服务中的一个普遍的目标是最小化能量消耗。WSN 服务的例子是数据聚集，处理和通信。网络管理服务层定义了一系列为执行特定的功能必须实施的操作/处理。其他方面是为这些功能定义范围、情况和参数。随着这些功能的执行，一个给定的服务便得到了保证。正如用户所理解的，管理功能表示了一个管理服务的详细的粒度功能分配。因此，管理框架必须包括 WSN 应用的功能设计图以及用户。

16.3.2.3　网络

这一层处理网络层要考虑到的事，例如节点协作、网络覆盖、连通性、数据聚合以及通信。传感器节点之间的关系被定义和维护。作为一个目标功能，网络增益在出现节点的感知，处理以及通信受限时得到优化。

16.3.2.4　网络元素管理

这一层覆盖了无线传感器节点的能量管理，移动管理，状态管理和任务管理。如果传感器节点在设计和功能上是异构的，管理系统必须考虑在设计一个 WSN 管理系统上的这些区别。

16.3.2.5　元素层管理

元素层管理处理那些需要被管理的或者执行一些管理功能的网络元素。基于一个应用的执行范围，一个管理元素可以是一个单一节点或者一组节点。当应用需要大量的节点时，网络元件与一簇节点相关。

16.4　设计管理结构的其他方案

一个管理结构的结构设计是其管理系统的核心。对于 WSN 的每个需求和目标，结构应该是轻量级的，可扩展的，以及自适应的。再者，它应该确保容错性，应答和低生产成本。在这一节中，我们讨论了普遍提出和采用的 WSN 的结构设计方法。

16.4.1　基于策略的方法

基于策略的管理是一种决定管理，执行和控制网络资源的运行规则的方法。策略定义了在不同的操作环境下部件的行为，这有助于提高适应性和网络的容错能力。定义了针对故障、结构、计费、性能以及安全管理的策略来提高网络的自治力和功效。例如，MANNA（A Management Architecture for WSN）[1]使用策略来定义网络管理者和代理之间的相互交流，为运行环境提供决策支持。

16.4.2　代表管理

WSN 是资源受限的网络，在这种环境下能量通常是最受限的资源。在一个受

限设备上注入太多的感知，处理或者通信代价将会耗尽它的能量。再者，与通信相关的成本通常远远超过了感知和处理代价。因此，将管理责任委派给欠约束的节点，并且聚集到达管理者的路径上的管理信息能够帮助节省网络能量和显著地延长网络寿命。这种情况下的管理责任是委派给靠近终端节点的节点，并且采用局部处理来延长网络的寿命。Guerilla management[6]和LoWPAN Network Management Protocol (LNMP)[7]采用这种方法，将管理分布到能量较高的节点上。

16.4.3 分布式管理

集中式管理经常是作为传统系统以及WSN中的一个执行瓶颈。集中式管理方法将通信聚集流向中心管理者导致网络中的活动链接拥挤。WSN流向中心管理者的通信流会耗尽在那个特定路由上的节点的能量。再者，由于网络划分的问题，网络的子部件有时可能与网络分离，集中式管理在WSN中是不受欢迎的。因此，分布式管理能够用来保证可靠性和能量有效性。此外，在分布式管理中，每个管理员有它们自己的领域，能够保证管理信息的局部处理，有助于降低网络延迟和网络通信量。这些管理员能够彼此通信，共享管理信息，这样能够保证更加可靠。MANNA[1]使用分布式管理，根据应用目标为不同的位置选择管理员和代理。

16.4.4 层次管理

在一个分层的方法中，管理任务被划分到不同的层中，每层独立地执行它的任务，如果需要的话提供与其他层通信的口。不同层的管理任务可以同时存在不同的模式。再者，任务的修改更加简单，这样有助于维护。在参考文献[13]中，作者采用一个基于分簇的中间件框架，使用分层方法，划分簇之间以及资源管理层之间的任务。在参考文献[14]中，作者使用分层的方法将任务分发到节点，网络以及特定应用层之中。

16.4.5 基于分层的管理

复杂系统能够被划分为小的可管理的系统部件。在传统上，管理是基于层次的，正如16.4.2节所提及的。对于每一层的管理能够定义不同的方法。例如，对于任务和服务管理，面向服务的结构（SOA）能够用于设计一个大的系统，将其划分为小的可管理的部件，并且每个服务作为一个WEB服务是可用的。更低层，也就是说，网络和网络元素层可以作为结合起来提供这些服务的功能单元而存在。MANNA使用这样一种方法，较低层次的信息作为功能单元被报告，较高层次服务位于使用这些功能单元的抽象层。Bridge of the SensorS (BOSS)[11]使用这样一种结构，管理信息是从传感器网络中收集的，作为用户可发现的可用服务。

16.4.6 移动或者智能的基于代理的方法

在网络管理中移动的基于代理的方法中，移动代理的处理任务被转移到管理数

据出现的位置。代码是本地执行的，仅仅将结果返回给管理者。这种本地数据处理结果消耗更少的网络带宽。再者，在网络划分的情况下，数据是由本地代理处理的，结果能够在连接恢复时发送，保证了可靠性。Agilla[15]是这种结构的一个例子，中间件帮助移动代理的部署。在参考文献［16］中提出了 WSN 的移动的基于代理的管理策略。在这种结构中，管理代理加强了操作策略和规则的执行。

16.5 已有的研究成果

16.5.1 MANNA

MANNA[1]是 WSN 管理中的一个基于策略的结构。MANNA 提出了在 WSN 中如何执行管理的技术基础。它提出了一个 WSN 管理的一个一般结构，强调它固有的对于开发它的应用的依赖性。它扩展了传统的二维管理，也就是说，“管理层”和“管理功能域”，（通过）添加“WSN 功能”作为一个新的维来使用 WSN 被考虑过的网络功能。新的 WSN 功能维覆盖了 WSN 的结构、维护、感知、处理以及通信问题。

这里，管理信息映射为 WSN 模型。基于由这些模型所获得的信息，管理功能和服务根据应用指定的策略执行。MANNA 描述了许多不同的网络模型，例如，感知和通信覆盖域映射，行为和依赖模型，网络拓扑和能量使用状态，以及设备成本。管理功能由从这些模型获取的信息组成，一些管理功能的例子是拓扑-映射-发现功能，能量-层-发现功能，以及覆盖-区域-监督功能。这些服务使用这些从管理模型获取的数据。最后，将这一系列功能结合起来根据为应用定义的策略来提供服务。

MANNA 结构被划分为三类：功能的、物理的和信息的。功能结构说明了不同的管理实体的管理责任的代表，也就是说管理员、代理以及管理信息库（Management Information Base，MIB）。它定义了在传感器网络中部署管理员和代理的最优位置。物理结构定义了如何执行功能结构。它定义了可以令每种应用类型满意的协议轮廓，并且提议不同的管理实体之间的接口应该采用轻量级的协议栈。信息结构提供了在网络中应该被支持的信息。这种模型根据信息的功能，将它们划分为不同类型的支持和管理类。被管理的类包括与网络、管理元素、设备、系统、现象、连接以及环境相关的属性。

16.5.1.1 MANNA 的 WSN 功能的方面

除了传统网络管理的两个方面，正如 16.3 节所提及的，考虑到 WSN 的特性，MANNA 结构提出了一个新的方面，WSN 功能。这覆盖了五个主要的 WSN 功能：配置、感知、通信、处理以及维护。这些如下：

16.5.1.1.1 配置

WSN可以由构成它的节点和网络的结构来描述。结构可以是同构的，网络中的所有节点有相同的硬件性能，或者它也能够异构的，网络由不同性能的节点组成。再者，拓扑和网络组织也会影响网络的构造。在一个基于树的拓扑中，节点能够分层地组织起来，或者网络能够有一个平面拓扑，所有的节点处于相同的层次。在元素层，传感器网络中的节点能够描述为sink节点，簇头节点或者源节点。

16.5.1.1.2 感知

它覆盖了传感器节点的数据聚集机制。数据收集可以是持续的，传感器节点持续地收集来自环境中的数据或者响应来自一个管理员的查询以及对环境中的一些事件发起数据收集。

16.5.1.1.3 通信

它覆盖了WSN的监听机制。监听可以是持续的，数据被周期性地发送到管理系统。它可以是基于命令的，回答用户的请求。它也可以是事件驱动的通信，传感器节点对一个感兴趣的事件发送应答信息。最后，它能够被编程，传感器节点根据应用定义的策略发送数据给管理员。

16.5.1.1.4 处理

处理可以是基本的处理，做简单的映射或者使用基于门限的过滤。反之，它能够是一个相互的关系，根据所收集的数据进行决策。

16.5.1.1.5 维护

这个功能描述了WSN中自动重构造，保护，优化以及修复的必要性。

16.5.2 BOSS

BOSS[11]定义了一个结构用于将Universal Plug and Play（UPnP）-based设备连接到non-UPnP-based传感器设备。UPnP允许设备的无缝连接，并且简化了网络部署。BOSS采用UPnP网络的自动设备和服务发现特性来进行WSN的服务发现管理。在这种结构中，一个服务代理监听来自传感器网络的服务广播。一旦接收，它便会使得服务通过UPnP是可用的。再者，BOSS也定义了一个UPnP设备管理服务，报告与设备信息相关的数据、环境、定位、同步化、能量、安全以及被发现的服务。因此，BOSS为WSN提供了一个网络管理和服务发现的结构。BOSS的优势是UPnP不需要进行配置，并且它能够轻易地与许多技术整合起来。不利条件是从传感器网络报告的数据量所产生的代价。

16.5.3 SNMS

传感器网络管理系统（Sensor Network Management System，SNMS）[17]为WSN提出一个应用合作管理系统。它由三个核心服务组成。首先，它包含有一个查询系统，能够快速地获取用户需要的关于网络健康和性能数据。第二，它包含一个登录

系统实现系统产生的时间的记录和检索。最后，它包含一个轻量级的网络层，能够与应用的网络层并行运行[17]。系统提供（支持）两种传输模式：收集和分发。收集需要用来获取来自网络的健康数据，分发需要用来分发管理命令和查询。对于收集模式，一个收集树协议定义为通过独立于邻居表来最小化状态需求，以及信息收集结束以后通过建立一棵树来最小化传输。建立的树是自适应的，可以根据网络拓扑进行调整。对于分发，协议保证了传输层信息分发的可靠性。它提出了 Drip，一种将所有信息可靠地分发到网络中的一个被管理设备的协议。每个部件希望使用 Drip 来登记一个特定的标识符，它代表一个可靠的分发信道。在那个信道上收到的数据将会直接交付给部件。它要求每个节点周期性地检查这个信道，并且隐藏从最近的在信道上提取的数据，作为对于周期性信息的响应，它要求它们返回隐藏的数据。这有助于即使当应用失败时也能够保证可靠性。SNMS 的主要优势是它的低的内存消耗以及低的通信负担。SNMS 也使用聚集技术打包查询应答来节省网络能量。

16.5.4 移动基于代理的管理策略

移动的基于代理的管理策略[16]引入了移动代理到 WSN 管理中来支持网络中的自治任务。对于 WSN 的移动的基于代理的策略管理，它们采用一个分等级的结构。这个结构是由策略管理作为较高层，簇策略代理作为中间层，本地策略代理作为较低层来组成的。一个策略管理员管理多个簇策略代理，它作为一个全局的策略决策点（PDP）。簇策略代理作为中间的 PDP，本地的策略代理作为策略加强点（PEP）。策略是分等级地分发的，从策略管理员到本地策略代理。任何节点能够动态的接管任何其他节点的功能来保证残存性。这个结构也定义了一些策略信息库，这种策略信息库是用一种类似于简单网络管理协议（SNMP）的数据结构来定义的。

16.6 作为一个整合技术的 IP-USN

基于因特网协议的无所不在的传感器网络（Internet Protocol-based ubiquitous sensor networks，IP-USN）是由 WSN 和 IP 网络合并形成的。但同时，这种网络是非常独特和有挑战性的，与其他无线技术相比这些网络是非常不同的（见图 16-1）。

	基于IP	私有
技术	IP-USN	MANET, WSN
连通范围	访问、边界、核心、因特网	访问

图 16-1 IP-USN 包含所有连接范围

对于这种网络管理的考虑大部分依赖于由它们的独特设计所导致的与众不同的特性。对于一个 USN 的管理结构必须考虑 USN 与 WSN 和移动自组织网络（MANET）的不同。图 16-2 说明了 WSN、MANET、USN 在结构、操作以及功能上的不同。结构上，一个 USN 主要是

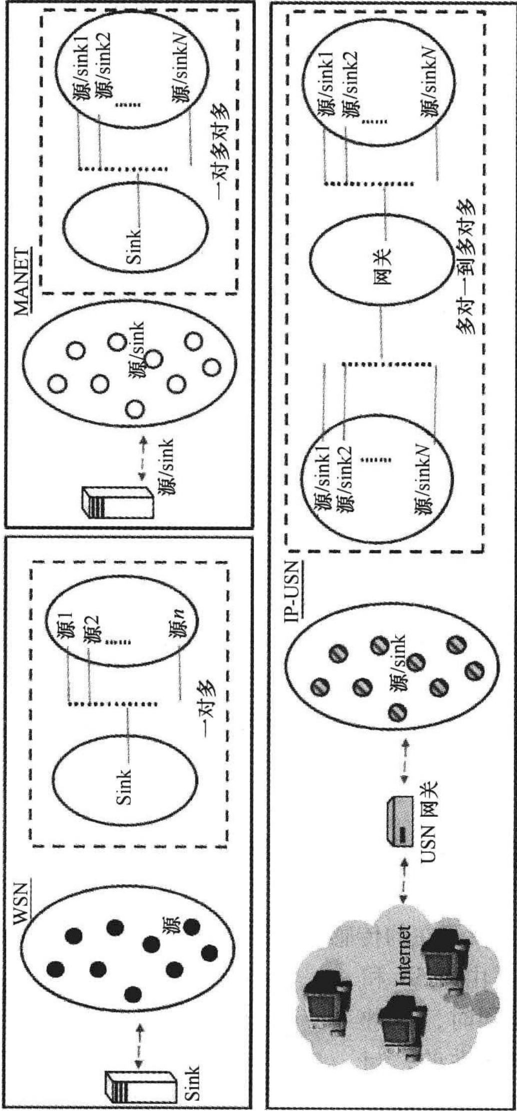


图16-2 WSN、MANET和IP-USN的架构简介

在网络实体和通信模型上不同于这些网络。操作上的不同在于 USN 的设备角色，特别是由于用户和应用异构。功能不同是由它们的特定应用的形成因素和代码足迹引起的。这些 IP-USN 的唯一特性对管理系统提出了具体的需求，见表 16-1。

表 16-1　对于 IP-USN 管理系统的考虑和要求

对于 IP-USN 需要考虑的	IP-USN 管理系统（IP-USN NMS）的要求
C. 1 IP-USN 显示出了用户异质性	R. 1 查询应该是跨用户域支持的
C. 2 通信是跨网络的	R. 2 网络管理框架必须认识和遵守网络信道行为
C. 3 网络部件多样和混杂的	R. 3 IP-USN NMS 的部件必须最佳分布于网络中
C. 4 网络间的句法和语义不同	R. 4 转换器和代理应该被嵌入 IP-USN NMS，无论是否需要
C. 5 网络之间查询类型和范围不同	R. 5 必须定义一致的查询类型和制定的管理信息库（MIB）

这些需求直接影响了一个 IP-USN 管理系统的设计，如图 16-3 所示。第一个需求表示了在不同的网络实体中，对不同类型和形式的查询的查询产生和处理的支持，例如，因特网领域中的设备、网关以及 WSN 领域内的设备。第二个需求指出网络实体应该是足够智能的，能够处理端到端的查询以及受限的响应传输，这里无线即有线和多跳路径的信道行为可能变化多样，因为中间设备可能不能够传递这些通信。第三个需求说明了通过最优的管理角色部署将网络智能分发的重要性，管理角色即管理员或者管理代理，它跨越了网络中的设备。因特网中的不同形式的设备因素，网关，和传感器节点引起第四个需求，也就是说，加入了代理和转换网关（无论何地需要）来将复杂的句法和庞大的查询请求映射为更轻的可管理的请求，同时保持相同的语义。最后一个需求表明了根据潜在的网络特性定义 MIB 的需要，来实现需求管理信息的信息获取。

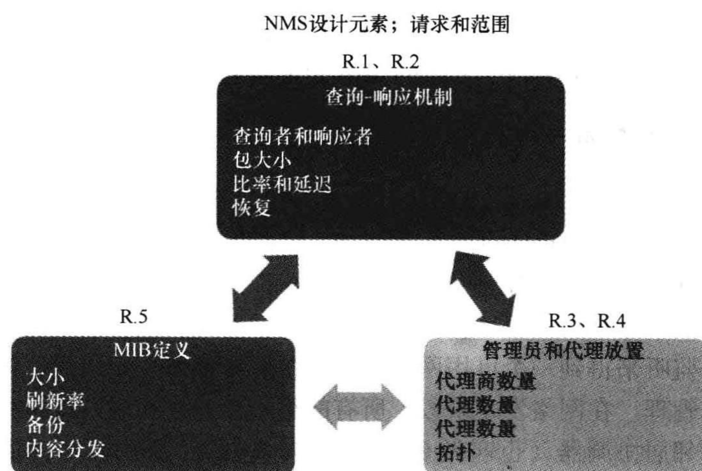


图 16-3　IP-USN 管理要求（覆盖层）与设计元素（嵌入）之间的关系

管理范围引导网络管理站（NMS）的目标，包括它的协议设计。这些形成一个基本的USN-NMS设计的目标，将在下一章介绍。

16.6.1 IP-USN NMS的目标

这里强调WSN与IP网络的正交性，概述了USN管理系统的目标。

G.1 USN管理系统必须提供IP中的遗留管理协议的反向兼容，如SNMP和它的变种。

G.1.1 必须能够识别那些能够容易地执行SNMP的网络元素。

G.1.2 不能够执行SNMP的网络元件必须通过一个SNMP语法分析器和一个代理连接到管理系统。

G.2 USN管理系统必须占用最少的USN通信。

G.2.1 必须存在一个资源发现机制来避免与不可用节点之间的无效通信。

G.2.2 必须存在一个机制将网络元素分为可用的、存活的、睡眠的或者死亡的。

G.2.3 必须存在一个分裂机制允许将一个大规模的管理包分为最少数量的片段。

G.3 管理系统必须最优地将管理员和代理放置到网络元素上。

G.3.1 必须存在一个机制识别管理员和被管理的代理的数量需求。

G.3.2 必须存在一个机制将一个管理员和一个被管理的角色分配给合适的网络元素。

G.4 管理系统必须分发和使用MIB来确保信息的可用性。

G.4.1 必须存在一个机制决定所有网络元素MIB的组成。

G.4.2 必须存在一个机制确保MIB的正确性和可用性。

G.4.3 必须存在一个机制提供对网络元素失效的适应性。

G.4.4 必须存在一个机制分发新的MIB定义或者相关网络元素的增量信息。

16.6.2 LNMP作为一个例子结构

提出了LNMP[7]通过6LoWPAN的特性的身份认证和考虑（低能耗个域网上的IPv6）来提供一个健壮的自适应的管理结构。6LoWPAN被认为是一个基于IP的传感器网络的实现，在它们上激活IP便于传感器的普遍存在。LNMP对这些网络围绕一个可使用的信息结构。管理操作分两步执行。首先，执行网络发现，通过分等级的设备状态监听来得到一个可用网络元素的快照。设备发现以后，第二步是对可用设备的实际管理。在网络发现阶段，所有的终端设备在初始阶段报告它们的状态给它们的双亲邻居协调器（也叫做6LoWPAN转发器），协调器最终通过它的祖先协调器将状态报告给网关。网络初始化结束以后，仅仅那些对下级节点状态的改变才会报告给网关。这个结构通过实现SNMP在设备上的使用，也支持设备监视。由

于 6LoWPAN 网络的可用带宽是非常有限的，在最坏的情况下，仅仅可以发送 33B 的用户数据报（UDP），在 6LoWPAN 网关上使用一个代理来降低 SNMP 包的控制代价。对到来的 SNMP 包进行语法分析，转换为轻量级的本地管理信息，发送到目标设备，应答传回网关的 SNMP，发送到管理员。由 MIB 组成的信息结构，它是设计来考虑这些网络特性的。信息结构关注于重用现存的标准 MIB，针对 6LoWPAN 网络的特性定义了一个信息库。

16.7 网络管理作为 FCAPS 模型：一个新视角

在 USN 中"普遍存在性"可能是最显著的现象，为 WSN 给出一个新的意义。它意味着传感器网络可能根据用户的特权自发建立和分解，并且可能（整个）与其他传感器网络交互，要么与邻近网络直连，要么通过中间媒体与遥远的传感器网络相互作用。因此，将经典的传感器网络定义从以应用为中心修改为以用户为中心，从静态改为移动的，从源-sink 节点改为源/sink 节点-sink节点/源模式。通过使用 USN 这一角色，每个传感器网络通过一个随时间变化的用户群管理系统与超过一个应用相关联。对于 WSN 广泛使用的 FCAPS 管理方面应用已经做了大量的细节调查研究。然而，当把 USN 看做是 WSN 的自然转换时，FCAPS 的执行粒度持有一个均等的扩展前景。在这一章，我们围绕这个"值—增加" FCAPS 管理视角，仅仅介绍了更新的方面（见图 16-4）。

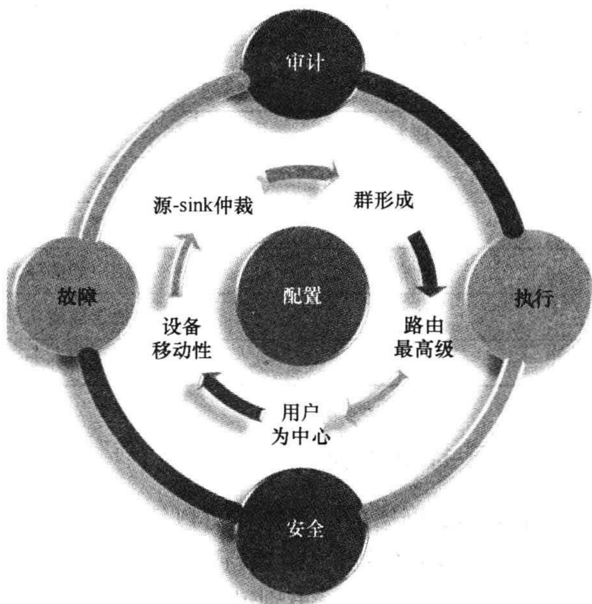


图 16-4 FCAPS 管理透视图

16.7.1 以用户为中心

在一个单一硬件平台上使用多种传感器使得 USN 比以前更加有用。USN 连接到异构网络，例如 IP 网络、电信网络和其他私有网络，给 USN 带来了大量的用户。每个用户可能发起一个不同的应用到同一个传感器网络。因此，传感器网络的应用特性变为用户指定的。USN 的用户为中心的特性发起了群形成和用户移动的现象。考虑到以用户为中心性，FCAPS 模型可能处理下面的管理问题（见表 16-2）。

表 16-2　不同管理域用户中心性

管理方面	用户中心性
错误	用户错误、设备错误、连接错误
配置	用户偏好、RAM 和设备占用
计费	访问控制、用户计费
执行	SLA 协议和违反 QoS
安全	用户认证、加密

16.7.2　群形成

在一个单一传感器网络上通过以一个逻辑群的形式抽象每个应用来附加地执行多个应用被证明是一个可行的提议，每个群可能有一个一对一的到一种传感器类型的通信。尽管每个群可能是彼此独立的，在所有的实际目的中，这些群可能依赖于一个共享的传感器节点集。因此，资源保留可能包含优先绑定，或者冲突解决可能对等地用于晚期绑定（见表 16-3）。

表 16-3　不同管理域的群形成

管理方面	群　形　成
错误	群失败、违背、资源饥饿
配置	群定义、节点竞争、选择、确认
计费	用户到群登录、每个群统计的节点
执行	群形成成功率
安全	群策略分支、群 AAA 管理

16.7.3　源-sink 节点仲裁

一个传感器网络的经典操作可以简单地描述为一个多源-单一-sink 节点模式。由于 USN 中采用多重群自由运行的概念，组织数据的节点和数据最终到达的节点可能位于同一个逻辑群，更别说同构的传感器网络。当一个传感器节点接收来自另外一个逻辑群的数据时，它可能被证明是一个逻辑群的源。这样的源和 sink 节点仲裁给出了路由处理复杂性的一个新的方面。因此，对于 FCAPS 管理是同样的（见表 16-4）。

表 16-4　不同管理域的源-sink 仲裁

管理方面	源-sink 仲裁
错误	源-sink 冲突、源超载、sink 超载
配置	发送和接收缓存大小、安排

（续）

管理方面	源-sink 仲裁
计费	活动源和 sink 号、网络寿命
执行	每个 sink 的有效吞吐量、每个源的吞吐量
安全	数据聚合和隐私管理

16.7.4　路由最高级

由于 USN 中的群形成的偏心率以及源和 sink 节点的结果仲裁，路由似乎成为占主导地位的活动。除非想出有效的方法来整理连通路由的最佳路由路径，USN 能量资源可能很快耗尽。因此，必须提出一个最优的路由结构通过识别任意源和目的地之间路由数据的空间和时间重叠来处理多余数据的传输。对这样一个路由机制和 FCAPS 之间的关系在表 16-5 中做了总结。

表 16-5　不同管理域的最高级路由机制

管理方面	路由最高级机制
错误	路由请求广播失败、路由错误
配置	路由核心设置、路由跳数
计费	广播号，活动的、过时的和无效的路由号
执行	每个路由的有效吞吐量、每个源的吞吐量
安全	路由核心广播合法性、黑洞尝试

16.7.5　设备移动性

在 USN 中，传感器节点可能被划分为静态的和可移动的。FCAPS 管理系统对于变化了网络动态的敏捷性强烈地依赖于它的诊断结果和对于传感器节点移动行为的响应。特别感兴趣的是一个管理系统抢占建立的中间设备移动性的空白的能力——如何迁移一个移动传感器节点上的活动时间和移动节点如何调整适应陌生的环境。由移动性引起的管理问题见表 16-6。

表 16-6　不同管理域的设备移动性

管理方面	设备移动性
错误	丢失的会话号、消失的设备号
配置	HA 和 FA 分配、会话迁移统计
计费	移动感知会话的 SLA、每个包的花费
执行	迁移成功率、迁移时间
安全	会话隐私管理、FA 的认证控制

16.8 结论

本章中，从广度上介绍了 WSN 网络管理。WSN 中的网络管理被证明和被认为是独一无二的，并且为设计考虑定义了特定的目标。然后介绍了管理范围，包括它的尺寸、功能和层次。介绍了当前的设计方法，然后是各种实际的实现，包括研究发起形式的和工业部署的。IP-USN 是 WSN 的扩展范围，作为对于 USN 的完整的技术来识别一整套的挑战，需求以及目标。然后介绍了作为 IP-USN 的一个例子的 LNMP。提出了 FCAPS 的较新演化方面来认证出现的现象，这将会决定将来的管理结构如何形成。

参考文献

1. L.B. Ruiz, *MANNA: A Management Architecture for Wireless Sensor Networks*, PhD. dissertation, Federal Univ. of Minas Gerais, Belo Horizonte, MG, Brazil, Dec. 2003.
2. L.B. Ruiz, I.G. Siqueira, L.B. e Oliveira, H.C. Wong, J.M.S. Nogueira, and A.A.F. Loureiro, Fault management in event-driven wireless sensor networks, in *Proc. ACM MSWiM Conf.*, Atlanta, GA, Oct. 2004.
3. L.B. Ruiz, F.B. Silva, T.R.M. Braga, J.M.S. Nogueira, and A.A.F. Loureiro, On impact of management in wireless sensor networks, in *Proc. IEEE/IFIP NOMS*, Seoul, Korea, Apr. 2004.
4. L.B. Ruiz, T.R.M. Braga, F.A. Silva, H.P. Assuncao, J.M.S. Nogueira, and A.A.F. Loureiro, On the design of a self-managed wireless sensor network, *IEEE Communications Magazine*, 43(8), 95–102, 2005.
5. W. Chen, N. Jain, and S. Singh, ANMP: Ad hoc network management protocol, *IEEE Journal on Selected Areas in Communications*, 17(8), 1506–1531, Aug. 1999.
6. C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo, An adaptive management architecture for ad hoc networks, *IEEE Communications Magazine*, 41(2), 108–115, Feb. 2003.
7. H. Mukhtar, S.A. Chaudhry, K. Kang-Myo, A.H. Akbar, K. Ki-Hyung, and S.W. Yoo, LNMP–Management architecture for IPv6 based low-power wireless personal area networks (6LoWPAN), in *Proc. IEEE/IFIP NOMS*, Salvador, Brazil, pp. 417–424, 2008.
8. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, Wireless sensor networks: A survey, *Computer Networks*, 38(4), 393–422, 2002.
9. N. Kushalnagar, G. Montenegro, G, J. Hui, J, and D. Culler, 6LoWPAN: Transmission of IPv6 packets over IEEE 802.15.4 Networks, RFC 4944, Sep. 2007.
10. M. Welsh and G. Mainland, Programming sensor networks using abstract regions, in *Proc. USENIX NSDI Conf.*, San Francisco, Mar. 2004.
11. H. Song, D. Kim, K. Lee, and J. Sung, Upnp-based sensor network management architecture, in *Proc. ICMU Conf.*, Osaka, Japan, Apr. 2005.
12. S. Meguerdichian, S. Slijepcevic. V. Karaya, and M. Potkonjak, Localized algorithms in wireless ad-hoc networks: Location discovery and sensor exposure. in *MobiHOC—Symposium on Mobile Ah Hoc Networking and Computing*, Long Beach,

CA, pp. 106–116, Oct. 2001.
13. Y. Yu, B. Krishnamachari, and V.K. Prasanna, Issues in designing middleware for wireless sensor networks, *IEEE Network Magazine Special Issue*, 18(1), 15–21, Jan. 2004.
14. Z. Li, X. Zhou, S. Li, G. Liu, and K. Du, *Issues of Wireless Sensor Network Management*, ICESS, LNCS, Springer, Berlin/Heidelberg, 2005.
15. C. Fok, G. Roman, and C. Lu, Mobile agent middleware for sensor networks: An application case study, in *Proc. IEEE ICDCS Conf.*, Columbus, Ohio, June 2005.
16. Z. Ying and X. Debao, Mobile agent-based policy management for wireless sensor networks, in *Proc. IEEE WCNM Conf.*, New Orleans, LA, Sep. 2005.
17. G. Tolle and D. Culler, Design of an application-cooperative management system for wireless sensor networks, in *Proc. EWSN*, Istanbul, Turkey, Feb. 2005.
18. W. Louis Lee, A. Datta, and R. Cardell-Oliver, WinMS: Wireless sensor network-management system, An adaptive policy-based management for wireless sensor networks, *Tech. Rep. UWA-CSSE-06-001*, The University of Western Australia, Perth, Australia, June 2006.

第 17 章　无线传感器网络中的部署

本章，我们将深入地讲解不同传感器网络中均匀部署和非均匀部署策略。大体上，观察区域由感知事件的地理不规则性来描述。我们提出传感器网络确定事件检测模型和概率事件检测模型，为 WSN 评估采用不同的公制。然后，我们统一地总结了 WSN 的部署问题，它可能是一个多目标问题，并且是 NP 难的。我们描述了文献中所有的精确的和启发式算法。我们执行所有的 WSN 部署策略，并且在四个方面作了比较：①部署传感器节点的数量；②事件检测概率；③网络连通性和④计算复杂度。

17.1　概述

WSN 是由一系列的小的资源受限的设备，这些设备由计算机处理器、内存设备以及传输和接收信息的无线信道组成。因此，WSN 能够观察和监测它们所处的环境。当监测到了感兴趣的事件，信息从一个节点路由到另一个节点，最终聚集到网关节点或者基站。最近，将传感器网络应用到广泛的应用中引起了众多科研工作者的热情关注，例如环境监测、军事目标跟踪、天气预报、家庭自动化和入侵检测等。

尽管传感器网络中的挑战是多种多样的，研究人员主要关注于基础的网络挑战，包括路由协议，最小化能量消耗，传感器节点定位以及数据聚集[1,2]。不幸的是，关于传感器部署处理方面的研究成果很少。

在本章中，我们深入地探讨在监测区域内，均匀和非均匀的事件分布的传感器部署的不同策略。本章的余下部分结构如下：17.2 节概述传感器监测模型，包括确定性的和概述监测模型。17.3 节介绍 WSN 评估的不同度量方法。17.4 节致力于描述文献中提出来的不同的传感器网络的部署策略，包括产生的问题形式，均匀部署策略和非均匀部署策略。最后，17.5 节做出总结，并提出 WSN 部署的开放性的问题。

17.2　事件监测模型

WSN 部署的问题的一个重要因素是传感器节点的监测能力。这种监测能力严重地依赖于物理和环境特性，例如障碍物、环境噪声、事件速度、传感器硬件可靠性、事件和传感器之间的距离等。因此，提出了大量的不同复杂度的理论监测模型。它们都基于应用需求，设备特性，以及环境特性。

一个传感器中可以执行两个主要的机制来决定一个事件是否出现在部署区域内。在第一种机制中，每个节点执行事件监测决定，独立于其他传感器节点来进行监测事件，这意味着无论事件是否发生，通信和传感器的合作者都不需要做决定。在第一种机制中，我们做一个本地的决定。与第二种机制相反，监测到事件的传感器节点之间的决定是协作式的。我们有一个分布式的协作决策。考虑到协作感知机制意味着传感器之间相互合作和转换对事件监测做决定。协作需要一个特定的数据-连接协议来实现一个分布式的事件监测决定。这种传感器之间的协作有许多优势，例如在事件监测中增加了监测概率，降低了能量消耗[3,4]，具有更少的不确定性以及错误警告。

大体上，本文中我们划分三种传感器监测家族。第一个家族是最简单的一个，称为比特监测模型。第二个是比较实用一个，称为距离-相关概率传感器监测模型。最后一个家族称为跟踪监测模型，是第二个模型的一个扩展，包含了每个点上的事件持续时间。

17.2.1　比特模型

比特监测模型被认为是最简单的事件监测模型。这个模型仅考虑了事件离传感器的距离，并且对其他物理环境参数进行了抽象。比特模型假设仅仅当一个事件出现于一个传感器的感知区域内这个事件才能够确定被监测到。换句话说，如果事件和传感器之间的距离小于一个感知范围，事件监测概率等于 1。否则，如果一个事件出现于感知范围以外，事件不会被监测到。大体上，一个感知范围半径 R_{max} 依赖于传感器的硬件特性。

如果我们假设一个传感器 s，位于坐标 (x,y)，一个事件 e，发生于坐标 (x',y')，那么事件监测概率 $P(s,e)$，　定义为

$$P(s,e) = \begin{cases} 1 & \| se \| \leqslant R_{max} \\ 0 & \text{其余} \end{cases} \tag{17-1}$$

式中，R_{max} 是感知范围；$\| se \|$ 是传感器 s 和事件 e 之间的欧氏距离 $e[\ \| se \| = \sqrt{(x-x')^2+(y-y')^2}\ ]$。

在描述区域覆盖问题时主要考虑这个模型，例如目标监测或者 k -覆盖问题[5,6]。这个模型的目标是简化分析，但是不幸的是，它并没有真正地反应传感器的感知能力。

17.2.2　概率监测模型

一个新的模型被提出，为了寻找更实际可行的，称为，距离-相关概率传感器监测模型，而不是一个比特事件监测模型。使用这个模型，事件监测概率与传感器与事件之间的距离成反比。更加精确地说，定义了亲密和最大化的传感器监测圆。

如果一个事件（事件①在图17-1中）发生在亲密圆内，那么认为事件监测概率是1。如果一个事件（事件②在图17-1中）发生在亲密圆外，但是在最大圆内，那么监测概率随着距离增大而降低。最后，当距离超过最大圆半径时，那么事件（事件③在图17-1中）不会被监测到。

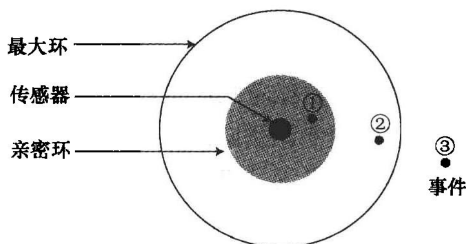


图17-1　传感器监测模型

在文献中，我们发现了根据传感器和事件之间的距离来决定总体监测概率的两个表达式。这些模型主要是由无线电信号传播理论来激发的[7]。如果我们假设一个传感器 s 位于坐标 (x,y)， 一个事件发生在坐标 (x',y')，那么这两个模型的事件监测概率 $P(s,e)$ 定义如下[8-10]

$$P(s,e)=\begin{cases}1 & \|se\|\leqslant 1\\ \dfrac{\alpha}{\|se\|^{\beta}} & 1\leqslant\|se\|\leqslant R_{max}\\ 0 & R_{max}<\|se\|\end{cases}\tag{17-2}$$

$$P(s,e)=\begin{cases}1 & \|se\|\leqslant 1\\ e^{-\alpha\|se\|} & 1\leqslant\|se\|\leqslant R_{max}\\ 0 & R_{max}<\|se\|\end{cases}\tag{17-3}$$

式中，R_{max} 是最大圆的半径；$\|se\|$ 是传感器 s 和事件 e 之间的欧氏距离；α、β 是传感器技术和事件特性参数。最加精确地，α 表示由于障碍物产生的能量失真因素。依赖于环境，β 通常小于5[11]。

17.2.3　跟踪监测模型

在参考文献［12］中，作者扩展了上面介绍的事件监测概率模型。这种新模型考虑到了事件持续时间。它的主要思想是当事件在一个点停留很长一段时间时增加事件监测概率。因此事件监测概率依赖于变量 t，t 用来量化一个特定点上的事件的持续时间。在这种情况下，事件监测概率是

$$P(s,e,t)=\begin{cases}P(s,e) & 0<t\leqslant T\\ 1-[1-p(s,e)]^{\lfloor t/T\rfloor} & t>T\end{cases}\tag{17-4}$$

式中，T 是感知算法执行以及事件发生时进行决策的时间段。这个模型能够用于跟踪应用。

17.3　部署标准

一定要记住的传感器网络需要做的是优化与应用需求相关的一个或者更多的度

量标准。以下，我们介绍最重要的度量标准。

17.3.1 部署传感器的数量

直到现在，主要工作是考虑低成本传感器。这个假设处理这个事实，嵌入到传感器内的感知设备假设监测非常简单的情形，例如温度、压力、湿度、光照、声音和磁性。然而，如果我们考虑部署成百上千的传感器节点，我们需要考虑到网络的全局成本。因此，部署传感器的数量是一个在 WSN 部署过程中需要处理的重要的度量标准。

另一方面，在一些特定的应用中，低成本传感器是不实际的。事实上，传感器成本高度取决于应用目标和监测环境。例如，海洋研究应用中，如污染监测，近海勘探，或者灾难预防，传感器成本是非常高的。这些环境中，感知部件能够更加精确地监测一些特殊现象并且能够抵抗环境影响。

17.3.2 覆盖和 k-覆盖

WSN 中的一个基础问题是覆盖问题。覆盖可以定义为部署区域内的事件至少会被一个传感器监测到的概率。这被称为 1-覆盖。显然，理想的部署中，覆盖率应该等于百分之百。不幸的是，传感器节点容易失效，发生测量错误，被毁坏，或者能量耗尽。这种情况下，定义了更普遍的 k-覆盖（$k \geq 1$），监测区域中的每个点至少能够被 k 个传感器监测到。

17.3.3 连通性

一个传感器的主要任务是监测它覆盖区域内的事件。然而，这个事件必须由传感器传输到一个特定的节点，称为 sink 节点，sink 节点负责将事件发生通知应用用户。如果 sink 节点在一个传感器节点的通信范围内，事件能够被轻易地传输到 sink 节点。然而，如果 sink 节点不在监测到事件的节点的通信范围内，事件通知必须以多跳的方式从一个传感器传输到另一个传感器，直到到达 sink 节点。在这种情况下，传感器网络拓扑必须是一个连通图，允许来自任何传感器的事件通知到达 sink 节点。

连通性和覆盖是相关的，因为它们都受到传感器位置的影响。在参考文献[6, 13]中，作者构想了充分的条件来保证 WSN 中的覆盖和连通。这个充分条件受到传感器位置，通信半径（R_c）以及感知范围（R_s）的影响。例如在参考文献[6]中，作者证明了，如果 $R_c \geq \sqrt{3}R_s$，并且区域是被完全覆盖的，那么通信图是连通的。

与覆盖规则相似，一些研究关注于 k-连通图，提出算法来建立每个传感器与 sink 节点之间的 K 不同路径。在这种情况中，我们的数据传输会更加可靠。如果一个路径失效，我们有 $K-1$ 个其他路径保持连通性。再者，建立 K 不同路径使得构想和实现分布式机制，如传输负载均匀，变得更加简单。因此，我们能够降低能量

消耗和延长WSN寿命。

17.3.4 检测概率

在17.2.2节中，当一个事件发生在覆盖圆内时它的检测概率并不等于1。为了测量事件检测概率，我们普遍使用式（17-2）和式（17-3）描述的模型。在真实部署中，每个点要求的事件检测概率可以小于1。它依赖于许多参数：区域重要性、事件发生频率以及事件传播增值。

再者，在许多传感器网络应用中（例如火灾检测警报、水质量监测等），监测区域能够根据事件的位置要求不同的监测级别。例如，在一个火灾检测系统中，危险区域（例如那些靠近化学物品存放位置的）要求具有高检测概率（接近于1）。然而，对于低火灾风险的地方，相对低的监测概率就足够了。

使用检测概率作为一个度量标准，我们能够计算部署满意率。这个度量标准代表了产生检测概率大于要求的检测概率的部署区域的百分比。换句话说，满意率代表了要求的检测概率分布和产生的检测概率分布的相似度。理想情况下，这两个分布应该是一致的。在实际部署中，很难获得一致的分布。

17.3.5 网络生命周期

一个传感器装有存储能量的电池来支持不同传感器硬件设备。电池容量很小，因此WSN的寿命严重地取决于能量消耗。为了保证一个特定的WSN的持续工作时间，我们需要部署合适数量的传感器于合适的位置。需要冗余传感器来达到一个用户要求的寿命。

节点部署对于WSN寿命有直接影响。当我们部署少量传感器，我们会获得一个弱的寿命。但是当我们部署许多传感器，寿命不会自动增加。在WSN中，当传感器数量增加时传输规模增加。因此，能量消耗增加，寿命降低。因此，部署更多的传感器节点并没有解决这个问题。为了延长寿命，我们需要考虑将更多传感器放置到正确的位置，例如，sink节点位置，事件频率，以及最终在传感器中执行的建立到达sink节点的路径（路由协议）的任何特定的机制。

17.4 传感器网络部署策略

本节中，我们形式化了部署问题。然后，我们将描述文献中的不同的部署策略。

17.4.1 问题定义

我们考虑一个传感器区域，用A表示。为了降低这个问题的计算复杂度，这个区域中分散化的。我们假设A是一个正方形，一个边等于n个单位。一个单位定义为一个规范的物理距离（例如1或者10m）。为了简化，本章余下部分，我们将利

用它的 baricenter 点参考每个正方单元 A。换句话说，当我们说一个传感器位于点 $p(i,j) \in A$，那么这意味着传感器位于相应的正方单元的 baricenter。相似地，计算一个单元方形的事件检测概率要考虑它的 baricenter 的检测概率。最后，我们考虑到发生在一个单元方形内的任何事件被一个位于它的 baricenter 点的传感器以概率 1 被检测到。

在一个有效的监测中，事件检测形式化为一个概率检测模型。每个点 $p(i,j)$，在 A 中与一个要求的最小概率检测门限，用 $r(i,j)$ 表示相关。

理想情况下，一个好的 WSN 部署算法应该获取 $\forall p(i,j) \in A$，那个点的测量的检测概率大于 $r(i,j)$，传感器连通图 G 是连通的。在一个点 $p(i,j)$ 的检测概率由所有在它的邻近地区的传感器来估计，但是事件检测模型是不协作的。点 $p(i,j)$ 的检测概率用 $P(i,j)$ 表示，由监测区域内所有可用的传感器来评估为

$$P(i,j) = 1 - \prod_{(x,y) \in Grid} [1 - p((i,j),(x,y))]^{D(x,y)} \tag{17-5}$$

式中，$D(x,y)$ 表示部署二价变量。如果 $D(x,y)$ 等于 0，它意味着没有传感器节点部署于格点 $p(i,j)$。如果 $D(x,y)$ 等于 1，它意味着一个传感器部署在格点 $p(i,j)$。

显然，如果部署了足够大数量的传感器，可能满足这个目标：$P(i,j) \geqslant r(i,j)$，$\forall p(i,j) \in A$。否则，考虑到成本，传感器的数量也是一个关键度量标准。除了满足最小的检测概率门限的要求，在一个部署问题中的第二个目标是最小化传感器的数量。形式化的，对于下面的多目标优化问题需要找到一个 WSN 部署解决方案：

1）最小化部署传感器节点的数量来满足下面的两个限制（2 和 3）。

2）对每个点 $p(i,j) \in A$，最小化要求的检测概率门限 $r(i,j)$ 和部署后的结果检测概率 $P(i,j)$ 之间的差距。

3）保证网络连通性。

上面描述的多目标优化问题是一个 NP 难问题。解决空间的大小是有限的，但是非常大（等于 2^{n^2}）。为了解决这个优化问题，我们可以选择一个精确的方法，如 Branch&Bound。不幸的是，由于它的指数级的复杂度，我们不能够将它应用到大的区域中。第二选择是利用启发式算法来解决这个问题。偏偏这些方法不能够保证获取最优的解决方案。然而，它的主要优势是多项式级的复杂度和运行时间，这是合理的。

下面，我们发展了文献中的精确的指数级算法。我们将它们分为两个家族。第一个家族重组在所有部署区域内要求的检测概率是一致的算法。第二个家族重组了所有的要求的检测概率在所有的部署区域内是不一致的算法。

17.4.2　均匀部署策略

当监测事件在所有的部署区域内有一样的重要性时我们讨论均匀部署策略。那意味着，所有要求的检测概率，$r(i,j)$ 在每个点 $p(i,j)$ 是一致的。我们能够在文献

中发现两个均匀部署过程，分别叫做 Random 和 Regular。

17.4.2.1 均匀随机部署

随机部署由在部署区域随机播撒传感器节点来组成。传感器的位置遵守均匀分布。选择均匀分布是由监测事件的同构特性激发的。

随机部署的主要优势是它的可行性。当部署区域不可达时，我们不能够手动将传感器放置到部署区域中。因此，仅能够随机部署，使用例如飞机，来解决这个问题。典型的应用是火山监督，战地应用等。随机部署有许多缺点。为了保证完全覆盖和网络连通性，我们需要部署许多传感器，因此，网络是庞大的，部署代价是昂贵的。如果传感器的数量大，会增加数据传输和能量消耗，网络的寿命也是一个关键的挑战。

17.4.2.2 规则部署

在一个规则部署中，传感器位置遵守一个规则的结构放置，如三角形、正方形、多边形等。因为监测事件是同构的，一个规则的部署保证了所有部署区域具有相同的传感器密度。规则的部署保证了一个充分的覆盖和网络连通性。但是它要求在监测区域部署传感器是方便的。

许多文献研究了规则部署并将它与均匀随机部署作了比较。在参考文献[14]中，作者证明了在规则的决策部署中的监测质量（QoM）比统一随机部署要好。在不同的规则拓扑间进行了比较，例如三角形、方形和六边形。在参考文献[15]中提出了一个增加的部署过程。它的主要的思想是如果没有传感器放置在将要部署的感知覆盖圆内，那么在点 $p(x,y)$ 部署一个新的传感器。在参考文献[16]中，部署过程建立了一个六边形拓扑。他们显示，对于要求的最小数量的传感器，六边形拓扑具有比三角形和正方形拓扑更好的执行性能。

17.4.3 非均匀部署策略

非均匀部署策略包括计算需要部署的传感器的数量和位置，并且考虑到了事件特征。正如1.4.4节所述，事件监测在部署区域内不是同构的。因此，区域由感知事件的地理不规则来描述。这之后，我们说明和发展了文献中的不同的部署策略。

17.4.3.1 最佳解决方案

一个非均匀部署问题能够描述为如下：

$$\begin{cases} \min\left[\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}D(i,j)\right] \\ \forall i,j \in \{0,1,\ldots,(n-1)\}:P(i,j) \geqslant r(i,j) \end{cases} \tag{17-6}$$

式中， $D(i,j)$，$P(i,j)$ 以及 $r(i,j)$ 分别表示部署二价变量，产生点 (i,j) 的检测概率和要求的检测概率。

这个问题用式（17-6）表示是一个典型的比特整数编程，它被证明是 NP 难

的[16]。为了计算最佳解决方案，我们应用 Branch & Bound 算法[17]。它是一个发现各种最优化问题的最佳解决方案的一个普遍使用的算法，尤其是在离散和组合最佳问题中。Branch & Bound 是对访问一套可接受解决方案的一个智能处理。当确定成本（数学证明）大于当前解决方案时，大量的候选子集针将会被删除。

Branch & Bound 的主要问题是运行时间和复杂度，在我们的实例中它是指数级的。为了给出一个 Branch & Bound 运行时间的思想，我们发起了三个具有不同区域大小的模拟场景：5 ×5、6 ×6 和 7 ×7。在图 17-2 ~ 图 17-5 中显示了要求的检测概率。对于区域 5 ×5，最佳的解决方案是在 12.02s 后发现的。对于区域 6 ×6，最佳的解决方案在 1，771.93s 以后发现。最后对于最后的场景 7 × 7，Branch & Bound 在 332，49h 后才得到最佳策略。我们注意到空间解决方案的规模增加地非常快。对于第一个场景规模空间等于 2^{25}，对于第二个场景规模空间等于 2^{36}，最后，对于最后一个场景规模空间等于 2^{49}。

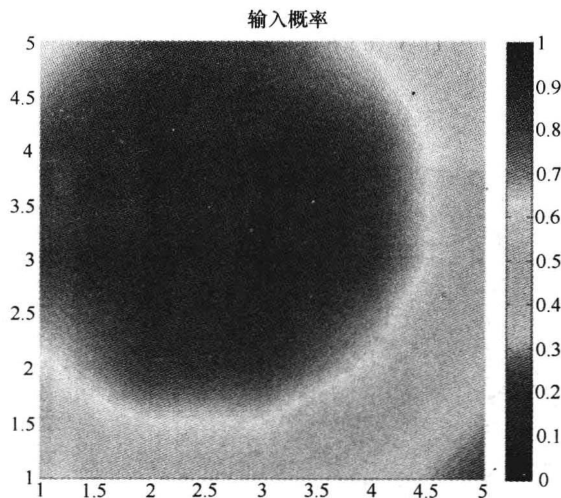


图 17-2 5 ×5 区域要求的监测概率

17.4.3.2 基于分布的随机的部署

在均匀事件中，基于分布的随机部署策略与随机部署策略相似。不同之处在于分布的选择。在均匀事件分布中，传感器的位置遵守均匀分布。然而，当一个事件不是均匀分布于部署区域上时，我们很可能需要部署更多的传感器在那些需要高检测概率的区域中。在这些限制下，我们需要选择一个部署后最大化满意度的传感器分布策略。

在均匀随机部署中的主要缺陷和优势存留于非均匀部署中。然而，在这种情况中，附加问题是传感器分布的选择。困难在于如何确定正确的传感器分布，而它应该与要求的概率分布一致。

在参考文献[18]中，作者研究了使用概率检测模型的随机部署过程。他们提出

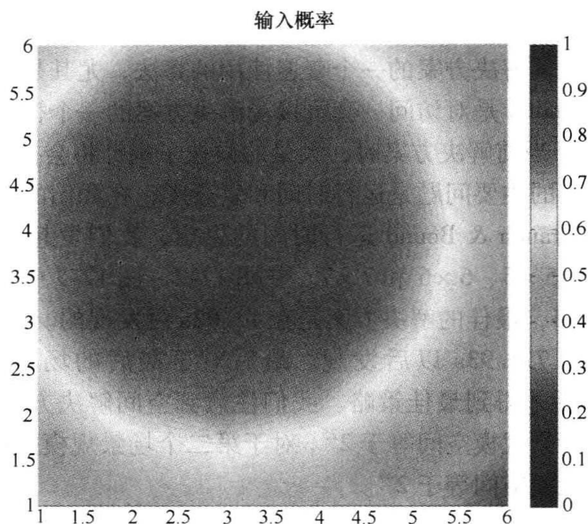


图 17-3　6×6 区域要求的监测概率

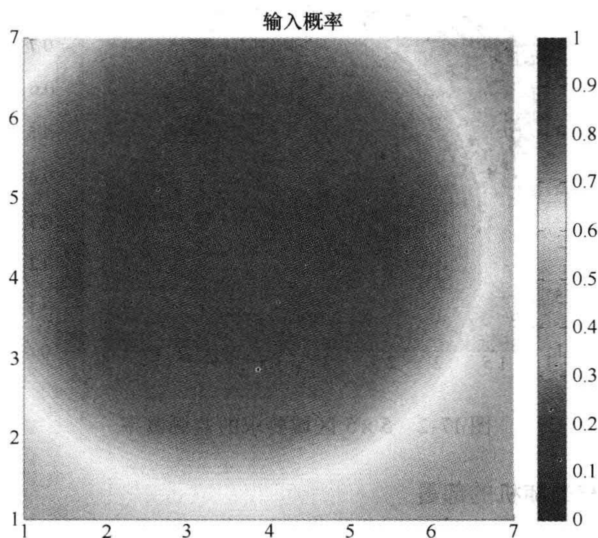


图 17-4　7×7 区域要求的监测概率

一个新的部署算法，叫做概率覆盖算法（PCA）。在参考文献[18]中，模拟结果显示 PCA 获取的事件检测概率比一个比特模型要好。在参考文献［19］中，在三个部署策略之间作了比较：①泊松过程；②均匀随机分布；③规则格。作者证明了一个格部署比随机部署渐近地回报更低的节点密度。因此，规则拓扑比同样性能的随机部署需要部署更少的传感器节点。

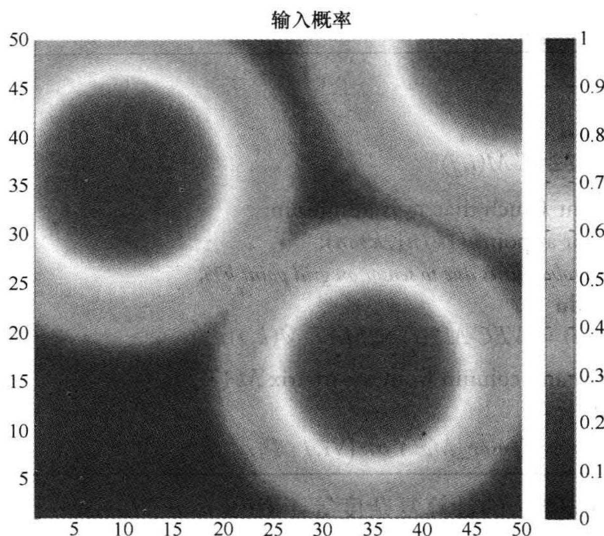


图 17-5　50 × 50 区域具有它要求的监测概率值

17.4.3.3　Max-Avg-Coverage

在参考文献[20]中，作者提出了一个部署算法称为 Max-Avg-Cov。这种部署策略的主要思想是最大化部署区域的平均覆盖。作者定义了三个主要的变量 $MAT-P$，$MAT-M$ 和 $VEC-M^*$。$MAT-P$ 是一个感知检测矩阵，$MAT-P=[P((i,j),(x,y))]$。我们在一个 $n\times n$ 的部署区域内评估它，整个格点等于 n^2。因此，$MAT-P$ 是一个方形矩阵，包含 n^2 行和 n^2 列。每个矩阵元素 $MAT-P(u,v)$，表示在点 v 被一个部署于点 u 的传感器检测到的概率。我们将坐标从二维转换到一维，因此，$u=y\times n+x$，$v=j\times n+i$。在 $MAT-P$ 中，当一个传感器被放置在点 $(\lfloor u/n\rfloor, u\%n)$，一行 u，在 MAT-P 决定了在部署网络的所有点内产生的检测概率。$MAT-M=[m_{ij}]$ 是一个遗漏概率矩阵。这个矩阵等价于 $1-MAP-P$。1 是所有元素都为 1 的方阵。换句话说，$MAT-M=[1-P((i,j),(x,y))]$。$VEC-M^*$ 是一个等价于 $\lfloor M_1,M_2,\cdots,M_{n^2}\rfloor$ 的向量。它决定了当部署过程结束时，每个点内的遗漏概率。开始时，没有部署传感器，因此，$VEC-M^*$ 初始化为全 1 向量，$VEC-M^*=[1,1,\cdots,1]$。

我们的目标是最小化需要部署的传感器的数量和最大化满意率。Max-Avg-Cov 是一个迭代算法；在每一步中部署一个传感器。Max-Avg-Cov 在 $\sum_v MAT-M(u,v)$ 最小的点 u 部署一个传感器，这意味着一个传感器部署在一个能够通过将格中的所有点遗漏概率的变化加和来最小化全局影响的点。当产生的检测概率大于要求的检测概率，或者部署的传感器数量超过了分配的传感器预算时，部署过程停止。Max-Avg-Cov 伪码在算法 17.1 中给出。

算法17.1: Max-Avg-Cov pseudo code

```
Num_Sensor = 1;
repeat
    for i=1 to n^2 do
        τ_i = Σ_j^{n^2} MAT-M(i,j);
    Select grid point k such that τ_k is minimum;
    Deploy a sensor at point (⌊k/n⌋, k%n);
    %Update miss probabilities due to sensor on grid point k%;
    for i:=1 to n^2 do
        VEC-M*(i) = VEC-M*(i) × MAT-M(k,i);
    Delete kth row and column from the matrix MAT-M;
    Num_Sensor++;
until (Num_Sensor > Sensor_Max) or (∀i,j : P(i,j) ≥ r(i,j));
```

Max－Avg－Cov算法的计算复杂度等于$O(Sensor_Max \times n^2)$。$Sensor_Max$的最小值是n^2。因此，计算复杂度等于$O(n^4)$。最后，这个过程建立了一个包含n^4个元素的矩阵，因此算法预留的内存大小是重要的。

17.4.3.4　Max-Min-Coverage

在参考文献[20]中，作者提出了另外一个部署算法，称为Max－Min－Cov。这个过程的主要思想是部署一个传感器在最小覆盖的格点。那意味着，我们将一个传感器部署于具有最高遗漏检测概率的点。

Max－Min－Cov和Max－Avg－Cov使用相同的数据结构。因此，$MAT-P$,$MAT-M$和$VEC-M^*$变量包含相同的数据。Max－Min－Cov伪码在算法17.2中给出。

Max－Min－Cov复杂度等于$O(n^4)$，它与算法Max－Avg－Cov的复杂度完全相同。再者，Max－Avg－Cov建立了一个有n^4个元素的矩阵，因此，当部署区域很大时，运行这个过程的内存规模是相当大的。

算法17.2: Max-Min-Cov pseudo code

```
Place first sensor randomly;
Num_Sensor = 1;
repeat
    %Update miss probabilities due to sensor on grid point k%;
    for i:=1 to n^2 do
        VEC-M*(i) = VEC-M*(i) × MAT-M(k,i);
    end
    Select grid point k such that VEC-M*(k) is maximum;
    Deploy a sensor at point (⌊k/n⌋, k%n);
    Delete kth row and column from the matrix MAT-M;
    Num_Sensor++;
until (Num_Sensor > Sensor_Max) or (∀i,j : P(i,j) ≥ r(i,j));
```

17.4.3.5　Min－Miss

在参考文献[21]中，作者提出了一个新的部署过程，称为Min－Miss。这个策

略是一个迭代算法，一个传感器部署在一个台阶。作者为每个格点定义了一个新的度量标准，称为超过遗漏概率，用 $\tilde{m}(x,y)$ 表示。后者量化了当一个新传感器被添加到点 (x,y) 时的覆盖增益。Min - Miss 的主要思想如下。首先，所有的可能空闲格点（这些点中未部署传感器）被选择来形成一个集合 AP。接着，对每个在 AP 中的点计算一个遗漏概率。最后，部署一个新的传感器到能够最小化超过遗漏概率的点，这意味着一个传感器被部署于能够最大化这个区域内的检测概率的位置。

Min - Miss 使用 $VEC-M^*$ 和 $MAT-OMP_{xy}$ 变量。$VEC-M^*$ 是一个与 Max - Avg - Cov 中定义的变量相同的变量。$MAT-OMP_{xy}$ 是一个与格点 (x,y) 相关的 $n\times n$ 方阵。它包含当一个传感器部署到点 (x,y) 时引入所有格点的遗漏概率。因此，$MAT-OMP_{xy}(i,j)=1-P((i,j),(x,y))$。$(x,y)$ 点的超过遗漏概率等于

$$\tilde{m}(x,y)=\sum_{i,j}MAT-OMP_{xy}(i,j) \tag{17-7}$$

Min - Miss 部署了一个新的传感器到点 (u,v)，最小化超过遗漏概率

$$\tilde{m}(u,v)=\min_{x,y}\sum_{i,j}MAT-OMP_{xy}(i,j) \tag{17-8}$$

当一个传感器被部署，Min - Miss 为 $VEC-M^*$ 更新遗漏概率。当所有要求的检测概率被满足或者所有可用传感器被部署，这个过程终止。算法 17.3 说明了 Min - Miss 的伪码。

算法 17.3: Min-Miss pseudo code

```
Num_Sensor = 1;
AP initialized with all grid deployment points, AP = {(x_i, y_i)}, i,j ∈ {1, 2, ..., n};
repeat
    %simulate sensor deployment in all points of AP%;
    for i:=1 to ||AP|| do
        Compute MAT-OMP_{x_i y_i};
        Compute over miss probability m~_{x_i y_i};
    Select a point (u, v) to deploy a new sensor;
    m~_uv = min_{x_i,y_j ∈ AP} (m~_{x_i y_i});
    Deploy a new sensor at point (u, v);
    %Update miss probabilities due to sensor on grid point (u, v)%;
    for i:=1 to n^2 do
        VEC-M*(i) = VEC-M*(i) * MAT-OMP_uv(⌊i/n⌋, i%n);
    %Update AP%;
    AP = AP \ {(u, v)};
    Num_Sensor++;
until (Num_Sensor > Sensor_Max) or (∀i,j : P_(i,j) ≥ r_(i,j));
```

我们注意到 Min - Miss 在部署中通过预测来决定将一个新的传感器部署到哪

里。这个预测通过模拟一个传感器部署到所有候选点（空闲的格点）来实现。然后，选择最好的解决策略。不幸的是，这种预测在运行时间上是昂贵的；计算复杂度等于 $O(n^6)$。如果我们选择最大化超过遗漏概率的解决策略，它包含了最坏覆盖情况下的部署。这种策略叫做 Max - Miss，在我们的情况下是没有引人注意的，因为它的唯一优势是给出了一个低带宽覆盖的思想。

作者也提出了一个新的检测概率模型。它是基于 17.2.2 节描述的模型。它包括通过任何算法计算出来的预先定义的位置在传感器部署中的不确定性。传感器位置中的错误可以用一个高斯概率分布来模拟。计算坐标 (x,y)，作为均值，x 和 y 方向的标准差分别为 σ_x 和 σ_y。一个传感器将被部署到点 (x_1,y_1) 的错误概率以及预先确定的坐标 (x,y) 等于

$$EP_{xy}(x_1,y_1)=\frac{\exp\left[-\frac{(x_1-x)^2}{2\sigma_x^2}-\frac{(y_1-y)^2}{2\sigma_y^2}\right]}{2\pi\sigma_x\sigma_y} \tag{17-9}$$

通过应用全概率定理，一个传感器应该被部署到点 (x,y)，以概率 P^* 检测发生在点 (x,y) 的事件：

$$P^*((i,j),(x,y))=\frac{\sum_{(x_1,y_1)\in A}[P((i,j),(x_1,y_1))EP_{xy}(x_1,y_1)]}{\sum_{(x_1,y_1)\in A}EP_{xy}(x_1,y_1)} \tag{17-10}$$

17.4.3.6　Diff-Deploy

在参考文献[12]中，作者提出了一个新的部署策略称为 Diff - Deploy。作者形式化了一个部署过程作为控制理论系统。一个部署矩阵 D（输入）被转化为矩阵 I（输出）。I 是定义为对数遗漏概率矩阵，$I(x,y)=\ln(VEC-M^*(y*n+x))$。我们有

$$VEC-M^*(i)=\prod_{x,y}(1-P((\lfloor i/n\rfloor,i\%n),(x,y)))^{D(x,y)} \tag{17-11}$$

对上面的等式我们应用一个对数函数：

$$I(i,j)=\ln(VEC-M^*(j*n+i))=\sum_{x,y}(D(x,y)\ln[1-P((i,j),(x,y))]) \tag{17-12}$$

已经证明这个系统是线性移不变（Linear Shift Invariant，LSI）系统。$g(x,y)$ 是描述一个输入 D 是如何转化为 I 的脉冲响应。$g(x,y)$ 定义为

$$g(x,y)=\ln[1-P((x,y),(0,0))] \tag{17-13}$$

如果我们使用式（17-2）描述的模型，那么

$$g(x,y) = \begin{cases} \infty & \sqrt{x^2+y^2} \leqslant 1 \\ \ln\left(1 - \dfrac{\alpha}{\sqrt{x^2+y^2}^{\beta}}\right) & 1 < \sqrt{x^2+y^2} \leqslant R_{max} \\ 0 & \text{其余} \end{cases} \quad (17\text{-}14)$$

因为这个系统是一个 LSI，输出系统 I 能够通过输入 D 的卷积结果和脉冲响应 g 来获得。式（17-5）显示了两维卷积

$$I(x,y) = D(x,y) * g(x,y) = \sum_i \sum_j (D(i,j) \times g(x-i,y-j)) \quad (17\text{-}15)$$

两维卷积可以通过建立特殊矩阵转化为矩阵乘积。这个技术在参考文献[22, 23]中进行了研究。最后，系统用如下表示

$$I_P = G_P D_P \quad (17\text{-}16)$$

式中，I_P 和 D_P 是有 n^2 个元素的向量；G_P 是一个有 n^4 个元素的方阵；I_P 和 G_P 分别代表要求的检测概率和检测模型使用的检测概率。目标是找到 D_P，它包含部署拓扑。因此，

$$D_P = G_p^{-1} I_P \quad (17\text{-}17)$$

在式（17-17）中计算 D_P 的元素与 0/1 不同。部署拓扑 D_P 并不是现实可行的，但是它给出一个放置传感器的思想。Diff - Deploy 主要的思想是部署一个新的传感器到与 D_P 最大值相应的格点。当一个新传感器部署以后，对 I_P 进行更新。重复同样的过程，直到满足了所有要求的检测概率或者部署了所有可用的传感器。Diff - Deploy 的伪码在算法 17.4 中说明。

算法17.4: Diff-Deploy pseudo code

```
%Compute the inverse of Gp%;
Gp^-1 = inverse(Gp);
remainIp = Ip ;
%Initialize all elements of Dp to zero (no sensors are deployed) %;
∀i : Dp(i) = 0;
repeat
    nextDp = Gp^-1.remainIp;
    Find the maximum value nextDp(k) in nextDp which satisfies the following
    constraints:

        ■ remainIp(k) < 0
        ■ Dp(k) == 0

    Dp(k) + +;
    Num_Sensor + +;
    %update remainIp %;
    remainIp = Ip − Gp.Dp;
    All positives values in remainIp receive zero;
until (Num_Sensor > Sensor_Max) or (∑_u^{n^2} remainIp(u) ≥ 0);
```

为了计算 Diff - Deploy 的复杂度，我们需要指定精确的算法来获取 G_P^{-1} 。文献中有许多反转矩阵算法，例如高斯消元法、LU -分解法、乔里斯基（Cholsky）分解法以及 QR -分解法。主要的不同在于复杂度。如果我们使用 LU -分解法算法，Diff - Deploy 的计算复杂度等于 $O\left(\frac{4}{3}n^6\right)$。因此，Diff - Deploy 是不可扩展的，并且它对内存的使用不是昂贵的（ $G_P c$ 包含 n^4 个元素）。

17.4.3.7 Mesh

在参考文献[24]中，作者提出了一个新的由图像处理和三维模型，也就是 mesh 表示，激发的部署算法。大体上，mesh 表示允许便利的任意表面建模，mesh 作为基本的原始事物来近似于一个表面[25]。在图像处理中，mesh 表示通过多角形元素用作一个图像的降低表示。有几种类型的 mesh。当所有的 mesh 有相同的形状（如三角形、长方形、六边形等）和相同的维度时，这种表示是规则的。不规则的 mesh 由异构的 mesh 形状组成。最后，使用一个特别的边大小不等的形状（例如三角形）可以获得分等级的 mesh 表示。

在无线传感器的部署环境中，mesh 节点表示传感器位置，并且每个弧线是两个传感器之间的欧氏距离。在区别对待的部署问题中适合用分层的 mesh 表示。应该考虑不同的形状，如三角形、长方形等。

提出的算法被称为区别部署算法（Differentiated Deployment Algorithm，DDA）。它的基本思想是只要认为是有益的，允许对 A 进行渐近地 mesh 划分。当代价函数 CF 减小（或者收益函数 PF 增加），mesh 划分就认为是有益的。否则，划分是不允许的，并且 mesh 是被标记的。

算法开始时考虑一个初始的未被标记的 mesh，它依赖于选择的形状。在长方形的情况中，通过四个角度划定部署表面来定义它。在三角形的情况中，初始未标记的 mesh 由两个三角形组成。表面的每个角至少是这两个初始未标记 mesh 的一个顶点。然后，算法评估代价函数 CF，获得当前的部署（传感器位于节点放置的位置）。然后，对于每个未标记的 mesh 执行一个划分测试。这里，关键的一点是选择新的传感器位置。有几种策略是可行的。例如，新的节点应该放到一个与划分的 mesh 的顶点等距的位置。另外一个选择是将新节点放到给定的 mesh 弧的中点。算法再一次评估代价函数 CF，获得新提出的 mesh 划分。如果前后两种划分结果代价函数值大于一个常数门限，这个 mesh 划分将被拒绝，并且作上标记。因此，DDA 能够接受一些不会减少代价函数来跳过本地极小值的划分。否则，提出的划分保留为一个下一步执行的候选列表中的可接受的操作。一旦测试所有未被标记的 mesh 划分结束，选择能够得到最小代价函数值的划分提议（候选列表中的）。推举的 mesh 是分开的，并且算法持续地检测新的 mesh 划分。当所有的 mesh 被标记或者所有传感器被部署以后，算法结束。

DDA 方法的伪码在算法 17.5 中说明。

算法 17.5: DDA pseudo code

```
Build initial meshes;
repeat
    Calculate the current cost function value, CF_a before division;
    Put the queue F to Ø;
    for Each mesh M not marked do
        Simulate the division of the mesh M;
        Calculate the new cost function value, CF_b, after division;
        if CF_b − CF_a ≤ Threshold then
            Put the mesh M in the queue F;
        else
            Mark the mesh M;
        if F ≠ Ø then
            Choose the best solution, M_best, in the queue F;
            Mark the mesh M_best;
            Accept division for the mesh M_best;
            Add the generated meshes;
            Num_Sensor ++;
until (Num_Sensor > Sensor_Max) or (All meshes are marked);
```

DDA 的计算复杂度等于 $O(n^6)$，因此运行时间随着未标记的 mesh 数量以及部署区域的规模增加。然而，DDA 内存使用不高。

17.4.3.8　分化的基于禁忌（Tabu）搜索方法的传感器部署

在参考文献[26,27]中，作者提出了基于元-启发式 Tabu 搜索的一个伪随机部署算法。Tabu 研究是一个本地搜索最优化技术，试图最小化一个代价函数 $F(x)$，x 表示一个参数向量，迭代地从一个策略 x 移动到一个邻近 x 的策略 x'（根据一个邻居函数 $H(x)$）直到满足了一个停止或者一个预确定的达到一个迭代次数 N。Tabu 搜索算法独立于检测模型。这个模型为这个方法提供了输入参数，尽管也可以使用一些其他的检测模型。

在 17.4.1 节描述了一个 Tabu 搜索算法用于应对分化的部署问题。下面详细介绍这个方法的初始化，邻居函数，代价函数以及新的指定阶数。

1. 初始化

Tabu 搜索的集合方法依赖于正确地选择初始化解决方案（s_0）。理想情况下，第一种解决方案必须是接近最优的解决方案；否则，因为迭代的最大次数是固定的，算法可能在达到最优解决方案以前就停止了。

作者认为部署一个传感器到一个点 $p(x,y)$ 的决策 $D(x,y)$ 是一个随机变量，它遵守一个参数为 $\alpha_{(x,y)}$ 的伯努利分布。决策规则的比特形式激发了选择伯努利规律。精确地说，$D(x,y)$ 能够假设值为 1 概率为 $D(x,y)$，值为 0 概率为$(1-\alpha_{(x,y)})$。

参数 $\alpha_{(x,y)}$，与一个点 $p(x,y)$ 有关，被选作为位于点 $p(x,y)$ 附近的点的百分比，并且它没有收到要求的检测概率。邻近点，用 $E_{(x,y)}$ 表示，定义为一套位于一

个将要被放置到点 $p(x,y)$ 的传感器最大监测圆内的邻近点。用公式表示为

$$\alpha_{(x,y)} = \frac{1}{\| E_{(x,y)} \|} \sum_{(i,j) \in E_{(x,y)}} 1_{\{r(i,j) > P(i,j)\}} \tag{17-18}$$

式中，$1_{\{cond\}}$ 是指示函数，如果状态 cond 为真它等于 1，否则为 0。Tabu 搜索方法的初始化阶段遵守这几步：

第一步：初始 Tabu 搜索解决方案开始时假设没有部署传感器。因此，$\forall p(x,y) \in A$，伯努利参数使用式（17-18）以及 $P_{(x,y)} = 0$ 计算。

第二步：产生一个列表，L_{init}，包括 A 的所有点。L_{init}是一个根据它们的伯努利参数逐渐减小的排序的点列表。

第三步：在 L_{init}中，选择有最大的伯努利参数的点 $p(x,y)$，并且将它从列表中移除。如果实际的检测概率 $P_{(x,y)}$，关联到 $p(x,y)$，低于 $r_{(x,y)}$，那么通过一个参数为 $\alpha_{(x,y)}$ 的伯努利决策规则随机产生一个策略来部署传感器到点 $p(x,y)$。

第四步：如果伯努利决策是部署一个传感器（$D(x,y) = 1$）），那么（1）对点 $p(x,y)$（集合 $E(x,y)$）附近的所有点的检测概率进行重计算，（2）更新（排序的）L_{init}。

第五步：如果 L_{init}不为空，回到第三步。

当满足第五步的停止规则，部署的传感器的结果位置被认为是初始的解决方案 s_0。后者保存在算法的内存，称为 Tabu 列表。本章的余下部分，我们将把这个列表称为 T。Tabu 列表的目的不是将方法限制在一个本地最小代价函数上。

2. 邻居探测函数

初始化阶段以后，一个 Tabu 搜索方法执行 N 次邻居探测阶段。这里，N 是一个选择的固定的参数，必须设置用来限制 Tabu 搜索迭代的次数。

在邻居探测阶段的第 n 次迭代中，产生并且评估了在前一次迭代中选择的这个解决方案的给定数量 V 的可能的邻居，标记为 s_{n-1}。邻居解决方案是可能的解决方案，能够通过一个基本转换从 s_{n-1} 达到。在 Tabu 列表 T 中提出的解决方案被认为是不可到达的邻居。

作者提出了两个邻居产生方法，称为面向抑制的阶段（Suppression-oriented stage）（H_{supp}）和面向附加的阶段（Additional-oriented stage）（H_{add}）。这两个方法在我们的 Tabu 搜索方法的连续迭代中交替来决定集合 V。下面具体介绍这两步。

面向抑制的阶段（H_{supp}）：这一步的目标是在那些多重覆盖的区域抑制一些传感器。这个方法的过程如下：

第一步：为所有的 $p(x,y) \in A$，使用式（17-9）并且假设在最后一次 Tabu 搜索迭代中获得部署计算伯努利参数。

$$\beta_{(x,y)} = \frac{1}{\| E_{(x,y)} \|} \sum_{(i,j) \in E_{(x,y)}} \left[\left(1 - \frac{r_{(i,j)}}{P_{(i,j)}}\right) \times 1_{\{r(i,j) < P(i,j)\}} \right. \tag{17-19}$$

第二步：产生一个列表，L_{supp}，包括所有的部署了传感器（$D(x,y)=1$））的 A 中的点。然后列表根据结果伯努利参数按降序排列。

第三步：在 L_{supp} 中，选择和删除有最高伯努利参数的点 $p(x,y)$，并且通过一个参数为 $\beta_{(x,y)}$ 的伯努利规则随机地产生决策来抑制在点 $p(x,y)$ 的传感器。

第四步：如果伯努利决策是抑制传感器（$D(x,y)=0$）），那么重新计算所有在点附近的点的检测概率。在这种情况下，使用伯努利参数与在列表 L_{supp} 中的每个点相关的新值来更新（排序的）L_{supp}。

第五步：如果 L_{supp} 不为空，回到步骤三。

一旦满足了第五步的停止规则，下一次迭代替换为一个附加阶段。

面向附加的阶段（H_{add}）：

这一阶段的目的是增加更多的传感器到实际部署中未充分覆盖的区域。执行类似于初始化阶段，除了第一步和第二步由下面的步骤代替：

第一步：对所有 $p(x,y)\in A$ 使用式（17-18），并且假设在最后一次 Tabu 搜索迭代中得到部署计算伯努利参数。

第二步：产生一个列表 L_{add}，包括所有 A 中未部署传感器的点。然后将列表一根据伯努利参数按降序排列。

三到五步，具体细节与初始化阶段相同，重复执行直到列表 L_{add} 为空。

3. 代价函数

第 n 次迭代中，邻居探测（一个基于抑制的或者一个基于附加的阶段）以后，在 V 个被探寻的候选者中选择一个解决方案 s_n。这个解决方案（不能够在 Tabu 列表 T 中）是一个由最小化一个给定的代价函数 F 来提供的。代价函数反映了在 17.4.1 节中描述的最优化问题的两个目标：最小化部署传感器的数量和最大化需要的检测概率的满意度。第一个目标可以通过计算部署的传感器的数量来量化。形式上，如 17.4.1 所定义的，如果一个传感器被部署到了点 $p(x,y)$，$D(x,y)=1$。否则，$D(x,y)=0$。为了最小化代价函数，F 包括下面的阶段：

$$\sum_{(x,y)\in A} D(x,y) \tag{17-20}$$

第二个目标通过下面的补偿函数整合到代价函数中

$$Penalty=\sum_{(x,y)\in A}\frac{[r_{(x,y)}-P_{(x,y)}]^{+}}{r_{(x,y)}} \tag{17-21}$$

式中，$[r_{(x,y)}-P_{(x,y)}]^{+}$ 表示在 R^{+} 中映射 $r_{(x,y)}-P_{(x,y)}$。形式上为

$$[r_{(x,y)}-P_{(x,y)}]^{+}=(r_{(x,y)}-P_{(x,y)})\times 1_{\{r_{(x,y)}>P_{(x,y)}\}} \tag{17-22}$$

根据补偿函数的表达式，一个好的部署解决方案应该能够获得比需要的检测门限更高（或者理想的，相等）的检测概率。如果不能够满足这点，补偿函数值转化为这种策略到达需求的门限之间的距离。这恰好是优化问题的第二个目标。

从上面的目标表达式，作者定义了两个代价函数，F_{supp}和 F_{add}。函数 F_{supp}用于在一个面向抑制的阶段中选择最好的下一次迭代解决方案。函数 F_{supp}使用式（17-21）来描述。另一方面，代价函数 F_{add}用于面向附加的阶段中。在这种情况中，与最优问题的每个目标相关的表达式通过下面的附加表达式整合为代价函数：

$$F_{add} = \sum_{(x,y)\in A} [D(x,y)] + Penalty \qquad (17\text{-}23)$$

在算法 17.6 中说明了 Tabu 搜索部署过程的伪码。

算法 17.6: Tabu search pseudo code

```
Compute initial solution s_0;
s_out = s_0;
bool sup-sensors = true;
Tabu-List = {s_0}, set of T last solutions visited;
for i=0 to N do
    neighborhood = Ø;
    if sup-sensors == true then
        for j=1 to V do
            s_i^j = H_supp(s_i);
            neighborhood = neighborhood + {s_i^j};
        Cost-Function = F_supp;
    else
        for j=1 to V do
            s_i^j = H_add(s_i);
            neighborhood = neighborhood + {s_i^j};
        Cost-Function = F_add;
    for j=1 to V do
        if s_i^j ∈ Tabu-List then
            neighborhood = neighborhood - {s_i^j};
    Select the best solution s_i^best in neighborhood, s_i^best minimizes the cost func-
    tion;
    Cost-Function(s_i^best) = min_{s_i^j ∈ neighborhood} [Cost-Function(s_i^j)];
    s_{i+1} = s_i^best;
    %update output solution s_out%;
    if Cost-Function(s_i^best) < Cost-Function(s_out) then
        s_out = s_i^best;
    %update Tabu-List%;
    Tabu-List = Tabu-List + {s_i^best};
    %Alternate between H_supp and H_add%;
    sup-sensors = not(sup-sensors);
```

·Tabu 搜索的计算复杂度等于 $O(NVm^2n^2)$。N 是 Tabu 搜索过程的迭代次数，V 是邻居的规模。这两个参数由设计者选择。根据特定的部署场景来校准它们。传感器

的一个覆盖圆覆盖了区域中的一系列的单元；m 代表一个传感器覆盖的子区域的长或者宽所占的单元格的数量。m 依赖于一个传感器的覆盖范围 R_c，并且 m 等于 $\left\lceil \frac{R_c}{\sqrt{2}} \right\rceil$。我们注意到，Tabu 搜索的复杂度依赖于部署区域的维度，传感器（R_c）的检测特征，和过程 (N,V) 的参数。如果 N、V 和 m 不大，积 NVm^2 等于常数 C_1，因此，我们能够说我们有一个二次复杂度 $O(C_1n^2)$。

在参考文献[28]中，作者扩展了在参考文献[27]中提出的 Tabu 搜索部署过程。扩展由保证网络连通性组成，因此包括了所有在 17.4.1 节中描述的最优问题（传感器数量，满足度，以及连通性）的限制。作者采用了参考文献[27]中提出的初始化阶段和在邻居中选择处理的解决策略来建立网络连通性。余下的阶段和代价函数与参考文献[27]一致。

17.4.4　部署策略对比

为了评估和比较上面提出的不同的部署策略的性能，我们实施了所有的方法，即，Random，Grid，MIN_MISS，MAX_MIN_COV，MAX_AVG_COV，Mesh-DAA，Diff-Deploy，以及 Tabu 搜索方法。采用 C++ 编程。

这个对比是基于不同的度量标准，如部署传感器数量，检测概率 θ（接收到的检测概率大于要求的检测概率门限的百分比）的满意率，计算复杂度，内存消耗，以及网络连通性。我们在两个阶段中比较部署方法。初始化阶段，我们关注于比较上面引入的所有度量标准，除了网络连通性。然后，我们关注于包括网络连通性的比较。

我们固定传感器的参数值 α,β,R_s 以及 R_c，分别为 1，1，5 和 3，选择小于 $\sqrt{3}R_s$ 的 R_c 来说明部署算法如何确保和建立连通性。我们记得如果 R_c 大于 $\sqrt{3}R_s$，它是保证连通性的充分条件，能够完全覆盖部署区域。我们考虑一个有 50 * 50 个单元的区域。要求的检测概率门限在图 17-5 中说明。

对于一个规则部署，我们选择一个网格拓扑，因此，形状是矩形。在网格方法中，我们选择三角形作为网格。为了划分一个网格，我们放置一个新的传感器到一个给定的弧的中间。代价函数是遗漏检测概率，等于 $(1-\theta)$。

我们通过设置迭代次数，Tabu 列表的大小，以及邻居规模相应地分别为 100，10，和 15，来调整 Tabu 搜索处理过程。对于随机部署，我们在 1500(10×15)个随机部署的拓扑中选择最好的部署拓扑。随机部署拓扑的数量等于所有由 Tabu 搜索策略产生的解决方案的数量。

图 17-6 显示了由 Tabu 搜索方法获得的部署传感器的位置。对于一个满意率 θ 等于百分之 96.52 的部署，部署传感器的数量等于 233。对于同样数量的传感器，当使用 Random，Grid，MIN_MISS，MAX_MIN_COV，MAX_AVG_COV，Diff-Deploy

以及 Mesh-DDA 方法获得的满意率分别等于百分之 77.12，83.40，85.52，83.96，82.56，96.2 和 93。

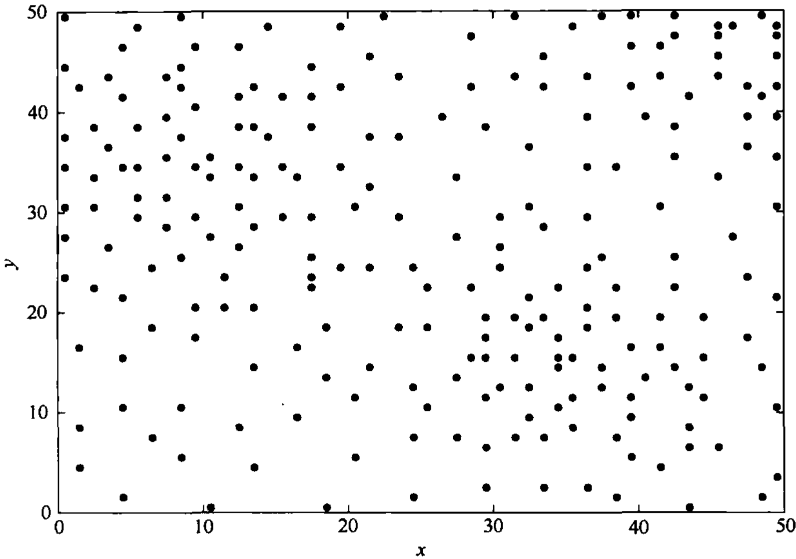


图 17-6 Tabu 查询方法的传感器位置

然而，为了达到由 Tabu 搜索方法获得的满意率，Random，Grid，MIN_MISS，MAX_MIN_COV，MAX_AVG_COV，Diff-Deploy 以及 Mesh-DDA 方法必须分别部署 450，552，337，373，518，234 以及 249 个传感器。

从上面的结果，我们能够注意到，Tabu 搜索和 Diff-Deploy 大大降低了部署节点的数量，同时也提高了满意率。但是 Diff-Deploy 的计算复杂度，等于 $O\left(\frac{4}{3}n^6\right)$，比 Tabu 搜索的计算复杂度 $O(c_1n^2)$ 大。还有，Diff-Deploy 的内存消耗比 Tabu 搜索中的内存损耗更加重要，因为 Diff-Deploy 操作有 n^4 个元素的方阵。Mesh-DDA 方法也给出了一个好的结果，但是它比 Tabu 搜索更加复杂。

在图 17-6 中，我们使用 Tabu 搜索方法策划了传感器的位置。与图 17-5 相比，我们能够注意到一个清晰的需要高检测概率门限的区域中部署的传感器（右上，右下，以及下右）。图 17-7 说明了由所有 X 的部署策略获取的累积分布函数(cdf)，X 是一个随机变量，在集合 $\{\vartheta_{i,j} \mid \forall i,j \in A\}$ 中获取值。形式上，

$$\vartheta_{i,j} = (r_{(i,j)} - p_{(i,j)}) \times 1_{\{r_{(i,j)} > p_{(i,j)}\}} \tag{17-24}$$

ϑ_i 取值为[0,1]。在图 17-7 中的每个弧为每个 δ 指明了 x 轴上的概率 $P(X \leqslant \delta)$。从上面的图片，我们能够注意到，Tabu 搜索、Diff-Deploy 以及 Mesh-DDA 方法比所有其他的部署算法提供了更好的性能。再者，我们能够观察到与其他方法相比 Tabu 搜索方法满意率很快达到区域单元的百分之百。这意味着未满足的单元接收到一个

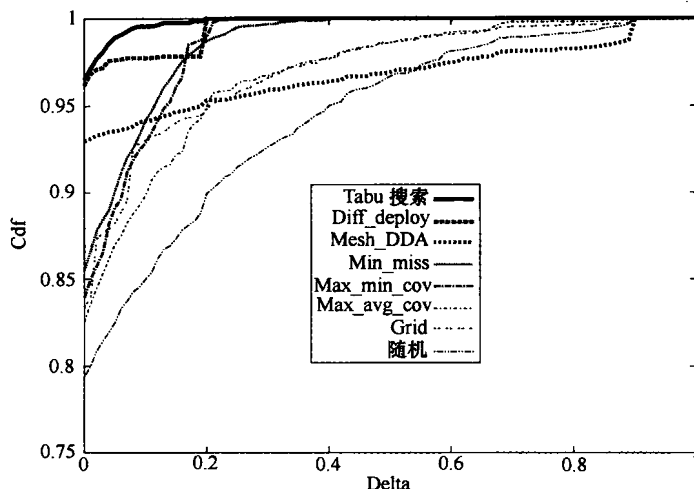


图 17-7　Cdf 要求的和获得的监测概率

非常接近于要求的检测概率门限的检测概率。

为了比较不同方法之间的网络连通性，我们运行第二个版本的 Tabu 搜索，它的连通性由结构来保证。Tabu 搜索方法部署 276 个传感器，满意率 θ 等于百分之 98.68。使用同样数量的传感器，Random、Grid、Min_Miss、Max_min_cov、Max_avg_cov、Diff_Deploy 以及 Mesh_DDA 方法的满意率分别为百分之 85.84、75.28、91.16、86.36、84.96、100 和 100。我们能够观察到 Diff-deploy 和 Mesh-DDA 的满意率达到了百分之百。图 17-8 显示在式（17-24）中定义的随机变量 X 的 cdf。

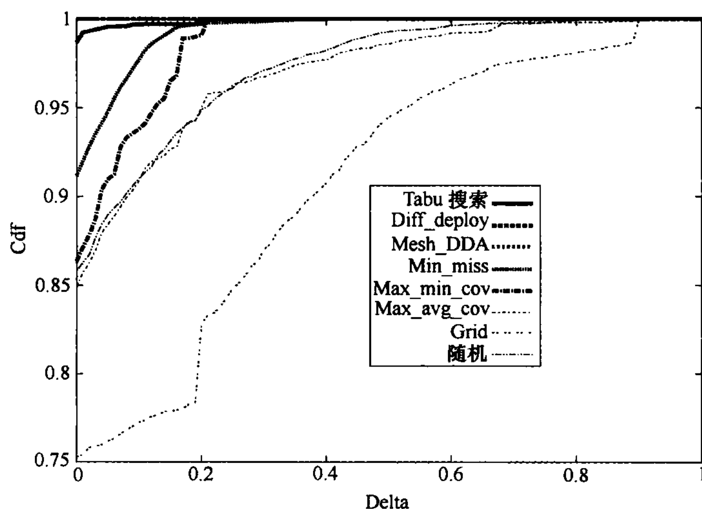


图 17-8　Cdf 考虑到连通性，要求的和获得的监测概率

我们计算由不同的部署策略产生的所有部署拓扑的连通部件的数量。如果连通性图片仅包含一个连通的部件，那么网络是连通的。图17-9和图17-10说明了Tabu搜索和Diff-Deploy的连通性。我们能够清楚地看到只有Tabu搜索提供了一个连通图。

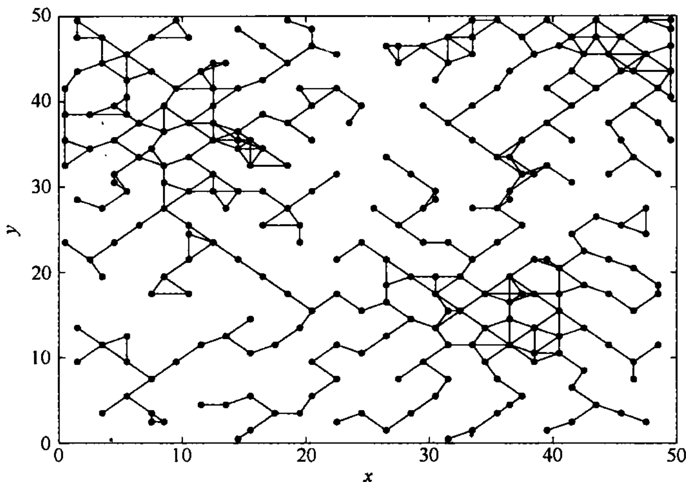


图17-9　Tabu搜索，连通图

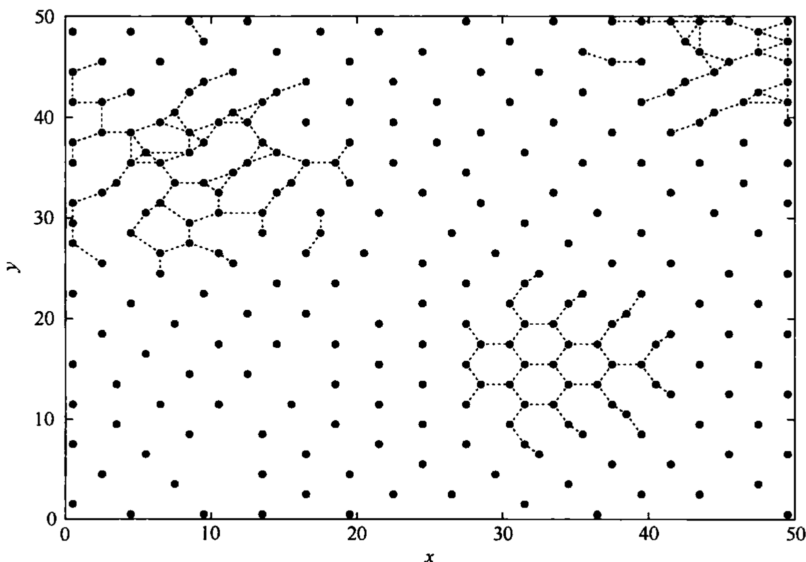


图17-10　Diff-Deploy，连通图

图17-11表示了由各种部署方法产生的与连通图相关的连通部件。我们给出了每个连通部件的节点的数量。我们仅仅划分了与每个图相关的10个最大的连通部

件。最大的连通部件是那个包含最大数量节点的部件。我们注意到，由 Tabu 搜索和 Grid approaches 产生的图片有一个连通的部件。在 Grid 中，我们确保连通性，但是检测概率的性能（见图 17-8）不够满意。部署在 Grid 中的传感器的数量略高于 Tabu 搜索。原因是由于网格形状的建立。其他的方法包含超过一个连通的部件，这意味着网络是不连通的。例如，Diff-deploy 和 Mesh-DDA 有 134 和 36 个连通的部件。Tabu 搜索方法是仅有的一种能够确保连通性的算法，并且与其他算法相比能够得到一个更好的满意率，同时最小化需要的传感器的数量。

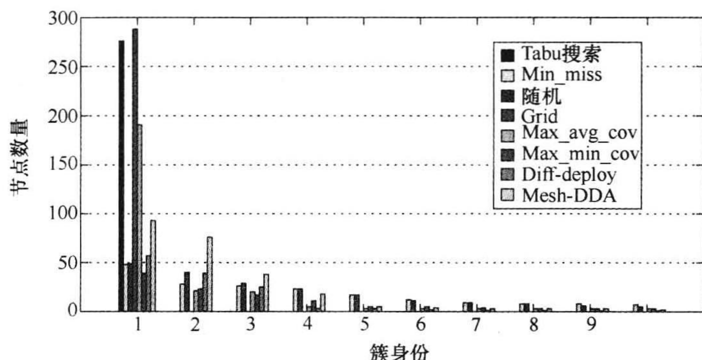


图 17-11 连通图的连通部件

17.5 结论和开放性的问题

这一章，我们深入探讨了传感器网络的部署策略。我们通过介绍文献中提到的主要的传感器监测模型来开始这一章。然后，我们强调了用于评估 WSN 部署的度量标准。之后，我们介绍了文献中主要的 WSN 部署算法。最后，我们提供了这些算法之间的一个执行性能比较。然而，由于版面稀少这一章中仍然有许多问题需要我们进行讨论，例如，能量消耗，依赖性，以及数据聚集。我们相信在这一章中讨论的部署问题可以作为建立一个 WSN 过程中重要的第一步。

参考文献

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, Wireless sensor networks: A survey, *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 38(4), 393–422, 15 March 2002.
2. J.N. Al-Karaki and A.E. Kamal, Routing techniques in wireless sensor networks: A survey, *IEEE Wireless Communications*, 11(6), 6–28, 2004.
3. P. Kumar, A. Mallikarjuna, and D. Janakiram, Distributed collaboration for event detection in wireless sensor networks, *Proceedings of the 3rd International Workshop*

on Middleware for Pervasive and Ad-Hoc Computing, Grenoble, France, pp. 1–8, 2005.

4. Z. Xue-Yu and C. Yang, Collaborative detection probability of mobile target for coverage in large-scale WSN, *International Conference on Wireless Communications Networking and Mobile Computing WiCom*, Shanghai, China, pp. 2710–2714, 2007.
5. C. Huang and Y. Tseng, The coverage problem in a wireless sensor network, *Book: Wireless Sensor Networks and Applications*, San Deigo, CA, pp. 115–121, 2003.
6. X. Bai, S. Kuma, D. Xua, Z. Yun, and T. Lai, Deploying wireless sensors to achieve both coverage and connectivity, *MobiHoc '06: Proceedings of the 7th ACM International Symposium on Mobile Ad-Hoc Networking and Computing*, Florence, Italy, pp. 131–142, 2006.
7. S. Kuo, Y. Tseng, F. Wu, and C. Lin, A probabilistic signal-strength-based evaluation methodology for sensor network deployment, *AINA '05: Proceedings of the 19th International Conference on Advanced Information Networking and Applications*, Taipei, Taiwan, 2005.
8. X. Li, P. Wan, Y. Wang, and O. Frieder, Coverage in wireless ad-hoc sensor networks, *ICC'02, IEEE International Conference on Communications*, New York, 2002.
9. C. Tai-Lin, R. Parameswaran, and S. Kewal K., Optimal sensor distribution for maximum exposure in a region with obstacles, *Global Telecommunications Conference IEEE GLOBECOM*, California, 2006.
10. A. Elfes, Occupancy grids: A stochastic spatial representation for active robot perception, *Book : Atonomous Mobile Robots : Perception, Mapping, and Navigation*, Vol. 1, pp. 60–70, IEEE Computer Society Press, 1991.
11. M. Hata, Empirical formula for propagation loss in land mobile radio services, *IEEE Transactions on Vehicular Technology*, 29(3), 317–325, 1980.
12. J. Zhang, T. Yan, and S.H. Son, Deployment strategies for differentiated detection in wireless sensor networks, *SECON '06: Proceedings of the 3th Annual IEEE Communication Society Conference on Sensor and Ad Hoc Communications and Networks*, Virginia, Reston, VA, Vol. 1, pp. 316–325, 2006.
13. Y. Wang, C. Hu, and Y. Tseng, Efficient deployment algorithms for ensuring coverage and connectivity of wireless sensor networks, *WICON '05: Proceedings of the First International Conference on Wireless Internet*, Budapest, Hungary, pp. 114–121, 2005.
14. S. Megerian, F. Koushanfar, G. Qu, G. Veltri, and M. Potkonjak, Exposure in wireless sensor networks: Theory and practical solutions, *Journal of Wireless Networks ACM Kluwer Academic Publishers*, 8(5), 443–454, 2002.
15. L.F.M. Vieira, M.A.M. Vieira, L.R. Beatriz, A.A.F. Loureiro, D.C. da Silva Jr., and A.O. Fernandes, Efficient incremental sensor network deployment algorithm, *SBRC Brazilian Symposium on Computer Networks*, Brazil, 2004.
16. M.M. Iqbal, I. Gondal, and Dooley L., Dynamic symmetrical topology models for pervasive sensor networks, *Proceedings of INMIC*, Lahore, Pakistan, 2004.
17. A. H. Land and A. G. Doig, An automatic method of solving discrete programming problems, *Econometrica*, 28(3), 497–520, 1960.

18. N. Ahmed, S. S. Kanhere, and S. Jha, Probabilistic coverage in wireless sensor networks, *Proceedings of the The IEEE Conference on Local Computer Networks 30th Anniversary*, Sydney, Australia, 2005.
19. Zhang H, and Hou J. C, Is deterministic deployment worse than random deployment for wireless sensor networks?, *INFOCOM*, Barcelona, Spain, 2006.
20. S. Dhillon and K. Chakrabarty, Sensor placement for effective coverage and surveillance in distributed sensor networks, *IEEE Wireless Communications and Networking Conference*, 3(20), 1609–1614, 2003.
21. Y. Zou and K. Chakrabarty, Uncertainty-aware and coverage-oriented deployment for sensor networks, *Journal of Parallel and Distributed Computing*, 64(7), 788–798, 2004.
22. H. C. Andrews and B. Hunt, *Digital Image Restoration*, Prentice Hall, Englewood Cliffs, NJ, 1977.
23. B. Hunt, A matrix theory proof of the discrete convolution theorem, *IEEE Transaction on Audio Electroacustic*, AU-19, 285–288, 1971.
24. N. Aitsaadi, N. Achir, K. Boussetta, and G. Pujolle, Differentiated underwater sensor network deployment, *IEEE/OES OCEANS'07*, Aberdeen, U.K., pp. 1–6, 2007.
25. L.B. Jordan, Progressive geometrical compression of arbitrary shaped video objects, PhD thesis, EPFL, 1998.
26. N. Aitsaadi, N. Achir, K. Boussetta, and G. Pujolle, Déploiement différencié des réseaux de capteurs, *8eme Colloque Francophone de Gestion de Reseaux et de Services GRES*, Hammamet, Tunisia, 2007.
27. N. Aitsaadi, N. Achir, K. Boussetta, and G. Pujolle, A tabu search approach for differentiated sensor network deployment, *Fifth IEEE Consumer Communications and Networking Conference IEEE CCNC*, Las Vegas, NV, pp. 163–167, 2008.
28. N. Aitsaadi, N. Achir, K. Boussetta, and G. Pujolle, Heuristic deployment to achieve both differentiated detection and connectivity in WSN, *IEEE 67th Vehicular Technology Conference—VTC–Spring*, Marina Bay, Singapore, pp. 123–127, 2008.

第3部分 RFID与WSN集成

第18章 RFID与无线传感器网络在架构和应用上的集成

射频识别（RFID）系统和无线传感器网络（WSN）是普及计算方面的两项关键的技术，由于在各自的应用领域中所带来的革命性，近几年已经引起了很大的关注。但是，这两项技术有着各自的研究和发展领域。

通过探索两种技术的优势，RFID系统和WSN的集成能够增加它们在其他科学与工程领域的使用。这一章节中，我们分析了RFID系统和WSN进行相互集成的重要原因，以及认识到实现高效率和有效一体化的关键要求。我们呈现出集成RFID和WSN的构架，提供了一系列现实中详细的应用例子。

18.1 概述

由于射频识别（RFID）系统和无线传感器网络（WSN）的重要优势和广泛应用性，在不久的将来，它们将是最重要的普及计算的新兴技术。RFID通信是快速的、便捷的，它的应用能够大幅度节省时间，提高服务，减少劳动成本，阻止假冒产品和盗窃，提高生产效益和保持质量标准。常见的应用范围包括高速公路收费，供应链管理，公共交通，控制建设准入，动物跟踪，开发智能家电等。

RFID系统主要用在没有提供任何有关物体物理环境的参考下识别物体或者对它们进行定位。另一方面，WSN是小型的网络，基于成本效益的设备，可以进行合作，通过感知进行收集并提供环境条件，如温度、光线、湿度、压力、振动和声音等信息。无线传感器网络提供具有成本效益的监测，包括工业控制，边境监测，环境关键应用监测，军事、家庭和医疗保健的应用。

RFID技术已经得到高度重视，并且已经在工业应用中进行了广泛地部署。另一方面，传感器网络一直是重大研究活动的焦点，但是除了军事上的应用之外它们主要是作为概念性的证据出现的。RFID和WSN的发展具有独立性的研究以及遵循发展道路，并导致了不同的技术。然而，也有许多应用中的一个对象的身份或地点是不够的，重要的额外信息可通过感知反演环境条件来实现。虽然传感器网络也可

以用于这些环境中，但是对象地点和身份的识别仍然可以通过 RFID 检索最重要的信息。在这些情况下的最优解是这两种技术的集成，因为它们是相辅相成的。

在这一章中，我们首先探讨 RFID 和 WSN 一体化的重要原因。然后，我们用一种有效灵活的方式列出集成 RFID 和无线传感器网络方面的关键性要求。随后，我们提出各种可能的集成架构以及讨论在实际应用中现有的集成方案和建议。对各种集成 RFID 和 WSN 的架构也进行可行性研究。这一章节中，我们提出以刘等人的提议为基础的不同的可能的综合架构。不过，我们进行详细而全面的研究，提出了各种新型的一体化商业战略与学术研究相结合的方法。

18.2 集成 RFID 和 WSN 的原因

RFID 和 WSN 技术的集成能够最大化有效性，是一个针对具有广泛可用性的应用提出的新观点，并且使现实世界和研究/学术世界联系在一起。因为由此产生的集成技术将提高扩展能力，可伸缩性和可移植性，以及减少不必要的成本。

1）能力和功能扩展：考虑到 RFID 网络可以提供诸如一个对象的身份和位置等重要信息，通过合并 RFID 和 WSN 的附加信息被检索，而利用这一信息的潜力是成倍增加的。例如，在供应链管理中，我们不仅可以跟踪食品，而且可以监督它们所在的环境条件以及检测食物何时腐烂。

2）可扩展性和可移植性：集成无线传感器网络的 RFID 系统享有无线通信的优势，与有线传输的不便和负担相比，传送和处理关键数据和信息是便利的，并且节省宝贵的时间。便携式 RFID 读写器，可进一步加速数据收集，简化不同应用能够的过程。例如，医疗应用包括对老年患者的监测诊断和对患者日常用药的监测，可以在完全不需要移动患者的情况下方便地收集繁琐的数据。

3）减少不必要的成本：在包括工业方面的许多应用中，减少就业服务的成本是一个关键因素。这项规定通过支持在非期望的情况下备份解决方案，以最小的成本达到预期的目标。例如，可以对易腐货物进行监测，假如它们没有被妥善保存其运输可被终止，从而避免不必要的额外的运输成本。

18.3 集成 RFID 网络和传感器网络的要求

RFID 和 WSN 一体化需要以这样一种方式展现，具体要求满足一个高效且有效地解决方法。一些需要考虑的最为重要的要求如下：

1）精确可靠的通信：在传统的客户服务网络中，大量的数据流从服务器向客户转移。然而，在集成 RFID 和 WSN 中，数据流主要从大量的客户机向一些服务器转移。随后，服务器将以一种可靠的方式处理所有从 RFID 和传感器收到信息，并在短期内采取适当的行动。在可以承受的等待时间内，数据所期望的可靠性和准

确性也将转交给集成系统的应用程序（或用户的）。集成射频识别传感器网络将数据安全地传送到目的地并提供一个成功完成任务的确认的能力是非常重要的。一个集成射频识别的传感器网络的可靠性和精确性也依赖于具有关键性的具体应用。在不重要的应用中，只需要低水平的可靠性即可。

2）能量高效：考虑到传感器节点和有源 RFID 标签是稀缺资源，集成射频识别传感器网络应该考虑到这个局限。集成系统应该是能量高效的保证准确和可靠的通信将在尽可能小的能耗下实现。

3）能量维持生存能力：由于大量的设备被应用于一个集成的射频识别传感器网络，对这样的网络最重要的要求是实现远程设备的配置和软件更新的能力。因此，我们可以以一个可接受的成本实现一个较高的生存能力和高效的网络维护。此外，重要的是在受到可能的拒绝服务器（DoS）攻击时，集成网络可以恢复。实现上述的方法是采用入侵容忍和减灾机制如复制关键网络设备。

18.4 RFID 和 WSN 一体化构架

18.4.1 集成 RFID 标签与传感器

集成传感器或者传感器标签的 RFID 标签，后续将做介绍，可以分为两个主要的种类：仅仅通过 RFID 读取器就能够通信的集成传感器标签和能够互相通信并且形成一个合作性的自组织网络的传感器标签。这一章，我们将要提出这两种集成传感器标签的主要特性，我们也将呈现出对有效研究的概述和针对这两种标签的商业目的。

18.4.1.1 通信能力受限的集成传感器标签

集成 RFID 网络和 WSN 的最简单的方法之一是在成 RFID 标签上集成感知能力。许多 RFID 标签在设计方面已经装上了传感器，因此，它们能够获取传感器的数据并传输给阅读器。然而，当 RFID 标签有感知能力时，RFID 网络和传感器网络之间的界限变得模糊，因为传感器标签用相同的协议和机制来读取标签 ID 和收集感应数据。在这种构架下，当集成传感器用来收集与环境、存在条件、相联系的对象（见图 18-1）的感知信息时，集成传感器标签作为一个普通的 RFID 标签起作用，它们赋予了一个唯一的标识。集成传感器节点的 RFID 标签通过 A/D 模块转换传感器的模拟信号，然后获得的数据由读卡器转发到基站[45]。

集成 RFID 传感器标签有许多商业和学术的价值，然而，在有源、半有源和无源集成 RFID 传感器标签之间是有区别的。

18.4.1.1.1 有源传感器标签

有源传感器标签使用电池给通信电路、传感器和微处理器供电。因此，它们有一个相当长的距离（大约 30m），能够实现较高的数据和传感器的活动率。然而，由于

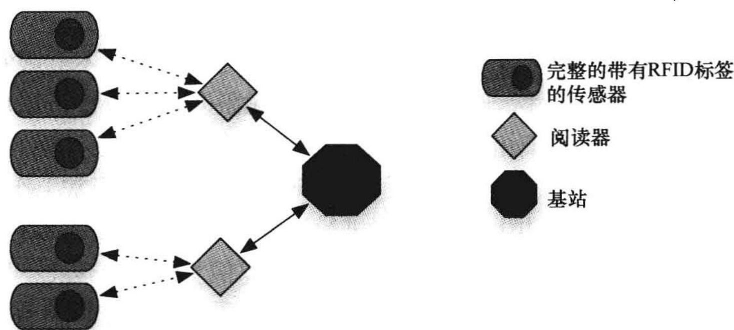


图 18-1　传感器标签、RFID 阅读器和基站的完整结构

使用了电池，设备的成本和重量增加的同时，RFID 传感器标签的寿命也受到了限制。

一种通过集成传感器与超高频 RFID 标签的 RFID 传感器标签和印刷的低成本环保纸张频率达到 950MHz 的有源 RFID 传感器标签是由费雷尔．维达尔等人提出来的[28]。提出的集成传感器标签使用嵌入式可充电电池薄膜，延长了电池的寿命。考虑到纸是最便宜的有机材料之一，提出的传感器标签呈现的优势很有吸引力，将触发大规模集成传感器的标签实现。日达等人[64]也提出了包含感知能力的，电池来源于低成本纸基材的 RFID 原型。

一种基于有源 RFID 的传感器嵌入式射频识别（SE-RFID）系统的设计是由邓等人[21]提出来的。这种系统的主要优势是传感器独立定期地采样外部数据，无论在该标签活动区域是否存在读卡器。邓等人[21]已经提出了关于 SE-RFID 的两种不同的构架。提出的第一种构架中，多个传感器可以嵌入到一个单独的 RFID 标签中，而在第二种构架中，每个传感器嵌入到一个单独的 RFID 标签。此外，他们已经通过开发一个使用 SE-RFID 技术的实时健康监测系统（HEMS）评估了所提构架的一种。在健康监测系统下，目标是开发一个不断监测的系统，这将能够不断监测，重新评估和诊断疾病。

一个有源振动传感器标签（24TAG02V）[11]和一个有源温度传感器标签（24TAG2T）[10]是应用比萨技术发明的。这两种传感器标签工作在 2.4GHz，它们的范围是 100m，都使用了防冲突机制。数以百计的标签可以同时被读，它们的电池可以使用四年。温度传感器标签从物体中收集实时温度并传输给读卡器进行登记。当温度到达一个特定的不能容忍的温度时，警报将会被触发。它的温度测量范围为 -50 ~ 150℃，精度为 1℃。振动传感器监测并记录连续或脉冲振动或冲击，它的最小灵敏度为 200mV/g，共振灵敏度为 4V/g，共振频率为 90Hz。

其他商业有源传感器标签包括 TELID310[39]（一个有源温度传感器标签由 Microsensys，Callistro 和 Elara 提出），有源温度和湿度传感器标签分别由 Adage Solutions[1]提出。

18.4.1.1.2 无源传感器标签

无源传感器标签从RFID读卡器获得工作能量。因此，电池不会限制它的寿命。它们有许多优点，例如小尺寸，低成本和长周期。无限寿命的特性可以应用于开发，无论是电池还是有线连接由于相对成本、重量或者其他原因都是（不）可行的。无源传感器标签主要的局限是它们必须靠近一个RFID读卡器才能正常工作。

一种集合了温度和光，并且可以用来进行环境监测的无源传感器标签被Cho等人[14]提出。这种传感器标签通过外部ISM带宽射频信号供能，它感知环境温度和光。Zhou和Wu[76]提出了一种嵌入式的无源超高频RFID传感器标签。他们的RFID系统包括一个900MHz的前端电路和基于镜子标准CMOS工艺磁传感器。它主要的优点是高灵敏度和低功耗。

一种长距离的无源RFID传感器网络标签的设计由Kitayoshi and Sawaya[43]提出。这种标签可以通信的距离超过10m，工作在2.45GHz和915MHz ISM带宽，它由一个分开的微带天线和无源电压倍增电路组成。为了证明这种标签的有效性，作者制作的无源温度传感器标签监测范围大于9m。

一种无线识别和感知平台的设计由Sample等人[66]提出。WISP是一种免电池的RFID传感器设备，与所有的无源RFID标签类似，通过RFID读卡器进行射频能量传输进行供电。WISP被制作成PCB板并且它的范围大约是4.5m。WISP是第一款集成微处理器的一种无源超高频RFID标签。

Instrumentel[31]提出了一种融集成无源传感器标签。在这种无源传感器标签中，通过电感耦合来供电，在没有电池的情况下可以激活标签。因此，传感器标签对于应用来说是理想的，而电池的重量和尺寸可能会干扰感知能力。这些无源传感器标签可以定制成多功能传感器，并且它们的接口需要通信协议。最重要的特性是它们含有板载微控制器、差分放大器以及多功能传感器。

微芯片[55]与Digital Angel合作已经设计和发明了一种无源嵌入式传感器标签，可以在不需要提醒糖尿病患者血糖水平的情况下，来测量动物和人身体的血糖水平。RFID传感器标签是无源的，通过扫描信号供电，不需要在传感器标签上添加电池。许多测量传感器需要一个精确地参考电压；为了实现它，使用的葡萄糖传感器有一个特殊的电路结构，对生理参数提供精确稳定的测量，来实现对葡萄糖浓度的精确测量。这项专利于2006年10月被授权并且命名为“嵌入式生物传感器系统”[55]。

一种能被用来测量动物体温的无源传感器标签也被Digital Angel提出来。这种称作Bio-thermo[9]的无源传感器标签是被嵌入到玻璃管形式的注射器，工作在134.2kHz的载波频率，遵循ISO 11785的标准[32]。它可以无创监测宠物温度，能够在早期阶段发现感染和疾病。然而，Bio-thermo没有可用的内存空间来存储数据。

商业上使用的集成传感器的无源RFID标签由三家日本公司发明，它们是OKI、NYK Logistics和HILLS[63]，以及Microsensys（TELID 210）[38]和Alien Technology（ALB-2484）[3]。

18.4.1.1.3 半有源传感器标签

当产生的射频功率满足操作时，半无源传感器标签作为无源 RFID 标签工作，否则它们工作于使用电池的半有源模式。Kim 等人[40]提出了一种集成无源和有源（电池供电）的半无源超高频 RFID 标签，支持 EPC Gen2 协议[25]。这种传感器标签当产生的射频功率满足操作要求时，作为一个无源 RFID 标签工作。在其他情况下，传感器标签工作于使用电池供电的半有源模式。传感器标签也使用一种可重复写的非易失性记忆库，它由铁电存储器（FeRAM）和片上温度传感器形成。

商业半有源传感器标签也已经被发布。德国 KSW-Microtec 公司已经生产了第一款集成了 VarioSens[44]传感器的半有源 RFID 传感器标签。VarioSens 传感器标签工作于 13.56MHz，符合 ISO15693 标准[33]。VarioSens 是一个可更新的有源传感器标签的版本，由称为 Tempsens 的 KSW-Microtec 公司生产。它有 1024B 内存，因此，可以在 292B 内可以容纳 720 个温度信息，而在 TempSens 中容纳 64 个数值。TempSens 仅支持密码保护，而 VarioSens 提供了三个等级的保护。VarioSens 中允许读、写、擦除数据的标签可以被定义。VariSens 提供的另外一个重要的特性是监测电池的水平并告知电池剩余的电量。它的工作温度变化范围为 -20 ~ 50℃，同时温度精度为 1℃。

Phase IV engineering Inc. 已经生产了一种叫做 SensIC RFID ASIC[56]的 CMOS 设备，它是能够测量和传输温度以及外部电容式 MEMS 传感器的感应值。SemsIC RFID ASIC 可能工作于无源和有源模式。它工作于 134.2kHz 的频率，并且符合 ISO 14223 和 ISO 11784/5 标准。它能够测量温度的范围为 -40 ~ 125℃，同时温度精度为 0.2℃。另一个叫做 ThermAssureRF 的商业半有源温度传感器标签是由 Evidencia[26]提出来的。

18.4.1.2 集成扩展通信能力的传感器标签

能够只与 RFID 读卡器通信的集成传感器标签被认为是 RFID 标签，有一些额外的感知能力和受限的通信能力。然而，用 RFID 标签集成传感器节点是可能的，以至于集成传感器标签将能够和其他无线设备互相通信。因此，这个种类包含的集成传感器标签超越了仅仅与 RFID 读卡器通信的局限，能够通过合作性的自组织网络互相通信（见图 18-2）。

Ruzzelli 等人[57]提出了集成 RFID 标签与传感器节点的研究方法。这个方法的主要目标是增加一个具有实时要求唤醒能力的传感器节点来减少能量的消耗，并消除传感器网络的空闲监听。这种方法称作 RFID 脉冲，通过把 RFID 标签赋予每个传感器节点来实现，也提供了 RFID 接收（读数）的能力。通过把 RFID 标签贴在传感器节点上，远处唤醒微处理器和发送接收传感器节点的需求是可能的。RFID 脉冲技术可以使用一个无源或者有源标签。

能够互相通信的集成传感器标签的一个典型例子是商业有效的 iRFID 标签[37]，它是一个有源的智能射频识别设备，由 Machine Talker，RFID 标签制造者生产，设

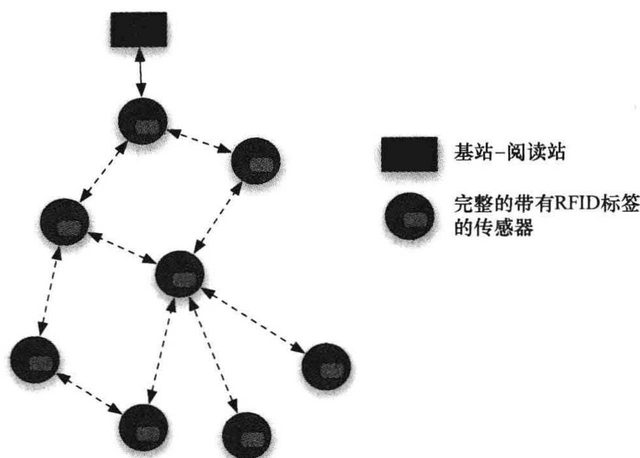


图18-2 完整的形成协作的自组织网络的传感器标签

计作为无线网络节点。iRFID标签是集成测量环境状况传感器的有源标签，诸如温度、光、振动甚至电池电量。这种标签工作于900MHz，通过一个专有的空中接口协议通信。当iRFID标签被激活并且它们互相接近时，它们能够自动形成一个无线mesh网络转移传感器数据。iRFID标签能够通信的范围是200m。根据这些设备的工作方式，它们可以通过WiFi或者有线网络协议与其他数据系统互相通信。iRFID标签已经在几家大型的炼油厂进行测试[8]。

能够互相通信的集成RFID传感器标签也被发明并应用于CoBIs方案[17]，这个欧洲的方案集中于商业进程和物质世界的管理。这种集成传感器标签可以收集、传送以及分享与周围环境相关的数据。它们的通信通过一个专有的点对点协议执行。每个标签对应一个移动传感器，这个无线收发器及其组件用于存储和处理相关业务管理规则。在传感器标签之间传送的数据包括它们唯一ID和环境中的感知信息。CoBIs RFID之间相互传送的距离是3m。在无线寻呼机中这个方案的试验通过在20～40个无线寻呼机的容器上放置集成CoBIs RFID标签来实现。根据使用的商业规则，集成传感器标签能够监督容器，并且当允许的特定地点存放的化学品数量受到侵犯或者潜在反应性化学物质存储的非常靠近时，就会触发警告。该集成传感器标签网络能够通过基站与更广泛的网络进行通信。

Sensitech也发布了一个叫做Temp Tale RF-enabled（TTRF）[68]的温度监测设备的RFID传感器集成设备。TTRF被建成一种有源RFID标签，它由温度传感器、射频芯片和天线组成。当有源标签传送感知数据给RFID读卡器时，传感器定时记录并存储温度。数据被集中收集起来，例如温度太高或太低，容易腐坏的物质有坏掉的危险时能够用来触发警报。这种集成传感器标签有电池供电和一个微控制器，可以在一个射频mesh网络环境中工作。它工作在915MHz或者868MHz ISM波段，可测温度变化范围为-30～70℃。

另一种类型的集成 RFID 传感器标签被 Aeroscout[2] 发明。更加精确地是，Aeroscout 已经生产了一种基于 WiFi 的有源 RFID 标签，应用于运动传感器，它有一个可选择的内部温度传感器，能够感知环境温度并且当达到设置的门限时触发警报。这种标签工作在 2.45GHz，为无线接入点（802.11 b/g）传送标准的 WiFi 信息。Aeroscout T3 标签（最新的版本）拥有 10 年寿命的电池，读取范围为 100m，1B 内存，可测温度为 0～100℃。

一种被加强型号的“多跳”标签由 NTT 实验室[62] 发明。这种标签不仅能够传送数据而且可以转发和读取数据。它们可被配置为“读卡器”或者转发设备。它们因电池供电工作在 429MHz 带宽，通信距离小于 1km。这种特别的标签最初是用来抵制猴子或者其他入侵者的，阻止它们肆意破坏农场和骚扰家养动物。为了实现这种标签，RFID 标签被贴在猴子或其他入侵者身上。当这些入侵者试图入侵农场时，它们会被 RFID 读卡器监测到，通过邮件提醒居民。同时，光或者声音警报可以用来恐吓入侵者。

18.4.2 集成无线传感器节点的 RFID 读卡器

另一种可行的与无线传感器网络相结合的 RFID 系统策略是通过集成 RFID 读卡器与传感器节点来实现的。在这种集成方案中，假定存在三种类型的设备：集成 RFID 读卡器/传感器节点，简单 RFID 标签和汇聚节点或者基站。这种一体化类型最初被张和王[75] 提出来。他们称这种集成 RFID 读卡器/传感器节点为“智能节点”。这种集成智能节点可以被看作传感器节点，能够用作 RFID 读卡器扩展它们的感知能力。智能节点能够转发信息并且可以被配置为无线传感器网络的转发节点。它们可以通过创建自组织网来相互通信。集成 RFID 读卡器/传感器节点可以作为路由器传送信息到正确的目的地。这些智能节点负责收集来自通信范围内的简单 RFID 标签的数据，通过互相通信来转发数据给汇聚节点/基站，在此所有的数据由一个人收集或者操作。这种集成网络的构架见图 18-3，类似于基于层次聚类的两层无线传感器网络。

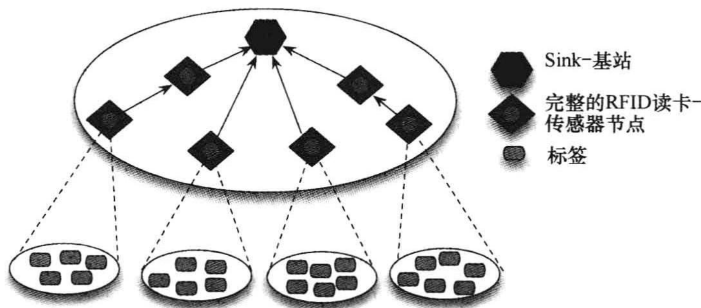


图 18-3 完整的带有无限传感器节点的 RFID 读卡器

这种类型的集成策略为可能的应用给出了新的观点。传统 RFID 读卡器仅包含无源操作，由于大体积造成的严重的可移动问题，以及天线的位置限制它们潜在的应用。集成智能节点不仅更小，更便宜，而且更容易部署。然而，这种集成策略也显示出一些严重的不足，因为它的特点是多到一的通信模式，并提出了一些相关的智能节点的能量不平衡问题。Yang 等人[74]认为，这种类型的一体化，由于智能节点有一个固定的传输范围，随着到基站的距离变近需要提交的通信量将要大幅度地增加。后来，那些距离基站节点更近的智能节点将会更早的用完电量，网络中部分区域将不再受到监控。Yang 等人[74]研究这种集成类型并且提出了一种策略，能够用来平衡网络的能量消耗并延长寿命。

Yang 等人关于平衡读卡器负载的提议基于在汇聚节点附近增加更多的读卡器。然而，通过在网络中增加更多的读卡器，网络的成本将会增加，会引起更多的冲突。这种劣势随着网络生存力的增加而增加。更加精确地是，Yang 等人[74]研究了有多少节点应该被加到汇聚节点的邻节点来获得最好的权衡。此外，他们提出了一种节点分布策略来实现能耗减少的均衡，使传感器节点的寿命最长。他们提出这种策略将大大地提高网络的寿命。

另一种集成 RFID 系统于传感器网络的方法由 Englund and Wallin[24] 提出。这种系统能够从一大片区域的 RFID 标签上收集数据。更加精确地说，他们已经（开始）关注系统的部署，在这里 RFID 标签可以克服传统 RFID 读卡器的距离限制读取远距离的数据。这是通过用射频收发器连接每个 RFID 读卡器实现的。所以，数据可以在相距 100 ~ 200m 范围内收发。因此，节点的一个完整网络被建立，能够作为路由器向正确的目的地传输信息。每个节点包含一个 RF 阅读器和一个射频收发器。为了使节点工作，微控制器用来协调每个节点不同的部件。更加精确的是，网络中的每个节点由微控制器、RF 收发器、RF 天线、RFID 读卡器、RFID 天线和电池组成。采用 Crossbow 技术的 Mica2 平台[20]用来部署这个方案。

集成 RFID 读卡器/传感器节点已经被商业化生产。SkyeRead M1-mini[69]是一种 RFID 读卡器，由 SkyeTek 生产。它的直径是 1in（1in = 0.0254m），厚度为 0.1in。这种读卡器的小尺寸使它适合应用于许多尺寸敏感的移动 RFID 设备中。M1-mini 的电池寿命能够延续超过两周的工作时间，提供每秒 20 标签的读速率。它工作在 13.56MHz 频率，能够读写 EPC 标签和智能化标签，以及符合 ISO 15693，ISO 14443 标准和 ISO 18000 标准的标签。此外，SkyeRead M1-Mini RFID 读卡器能够直接与 Crossbow Mica2Dot[48]传感器模块连接，产生一个集成 RFID 读卡器/传感器节点。另一个由 Alien Technology 发明的商业 RFID 读卡器是 ALR-9770[4]。它安装有四个天线设备，可以通过 802.11b/g 标准通信。

Gentag[27]提出了第三种集成传感器网络的 RFID 读卡器的解决方案。Gentag 是一家 IP 发展公司，为把传感器网络添加到移动电话、笔记本电脑、掌上电脑和其他无线设备的 RFID 读卡器上发布了一个专利。这种先进的专利是基础技术，使用

户可以使用手机来读取几乎所有类型的 RFID 标签。

18.4.3 混合结构

在混合结构中，RFID 标签和传感器节点物理上是不同的设备，但它们共存于一个集成的网络中独立工作。这种混合结构的主要优点是不需要设计一个硬件集成设备。然而，在 RFID 标签/读卡器和传感器节点之间可能存在通信干扰，因为在那种情况下，它们都是物理上不同的设备。为了避免这种干扰造成额外的费用，应该遵循这个过程。

最初，这种混合的结构被张和王[75]讨论过。他们认为，一种具有混合结构集成 RFID 传感器网络由三种设备组成：智能基站、标准 RFID 标签和标准传感器节点（见图 18-4）。智能基站是一种特殊的设备，它由 RFID 读卡器、微处理器和网络接口构成。智能基站没有能量限制，它们能够从 RFID 标签和传感器节点聚集信息，并把它们传输给本主机或者远处的网络。

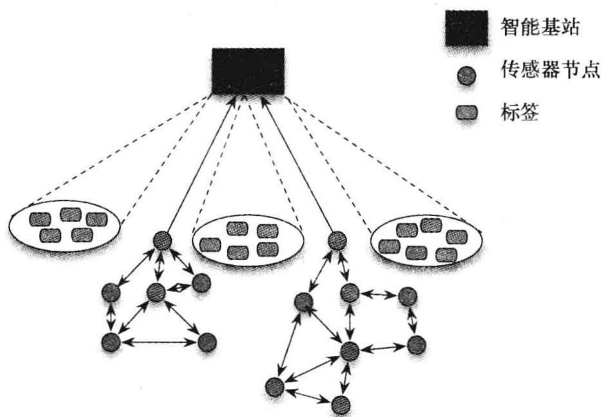


图 18-4　RFID 标签和传感器节点的混合架构

来自 RFID 标签和传感器节点的信息能够被传送到基站。因为智能基站没有能量限制，所以传统的网络协议结构也能重新部署。因此，智能节点不仅能够进行数据处理，也能够执行路由协议和传输协议，例如 TCP。能用在这种混杂的环境中的通信协议是 802.11/WiFi 技术。

这种具有混合结构的集成 RFID 和 WSN 的构架是 SARIF[13] 由 Cho 等人提出来的。根据这种构架，集成系统由集成服务器、RFID 网络和 WSN 组成。集成服务器是一个重要的组件，管理 WSN 和 RFID 网络的主要任务。RFID 网络由信息服务器、RFID 读卡器和标签组成，同时传感器网络由网关和传感器节点组成。RFID 网络的信息服务器与集成服务器进行通信，并传输与 RFID 相关的信息。集成服务器依靠从信息服务器获得的信息开始着手传感器网络中的任务。集成服务器也可能进入 RFID 网络并分配任务给它。作者也已经通过发展一种原型而对这种集成网络进行了评估，证明 SAFIR 通过负载均衡能够实现能量的高效。

一些根据混合结构能够支持 RFID 和 WSN 集成的商业解决方案已经被提出。例如，无论在什么地方，RFID 都是一个商业平台，包括丰富的特性、广阔的硬件、标准和支持协议，以及开发者和集成者需要来生产集成的 RFID/传感器设备的构架

灵活性。更加精确的是，无论在什么地方，RFID 都会允许高效的 RFID 标签/读卡器、环境传感器、条形码和移动设备的集成。

18.5 各种集成 RFID 和 WSN 的应用方案

18.5.1 医疗应用

RFID 和 WSN 已经独立的应用在医疗健康领域，诸如紧急治疗、轻微修复、移植牙齿以及跟踪医院的患者和全体员工。然而，集成这两种技术的潜力将会得到扩展。由于病人状况将能被持续监控并且如果病人的状况有突发的恶化，医生能够很容易的察觉，因此能够提高治疗质量。通过 RFID 可以跟踪病人的位置，同时他们的状况可以通过传感器监测。此外，通过使用集成 RFID 传感器网络[49]，诸如过时治疗、不精确的医疗水平以及高成本的医疗事故能够避免。在医疗领域集成 RFID 和 WSN 的可能应用包括感知温度、测量血压、心率或者 pH。在有关医疗应用的集成 RFID 和传感器网络的方法中，我们展示一些象征性的应用。

IMEC-Netherland（IMEC-NL）是一所荷兰的研究机构，已经通过使用集成传感器[71]的有源 RFID 标签建立了人类监测系统的原型。集成传感器标签用来记录和传送关于致命迹象的数据。人类监测系统主要用来调查诸如羊角风和无呼吸睡眠。直到现在，一个患有羊角风的住院病人通过电极监测，电极贴在他的脸上和头皮上，通过线连在一个盒子里。这个盒子搜集有关病人面部和大脑活动的数据信息。数据用来分析和跟踪病人的状况。类似的，患有无呼吸睡眠的病人通过贴在脸上的电极来测量眼睛、颚的肌肉运动以及大脑活动监测。IMEC 研究机构调查为病人提供可移动的可能性，甚至选择通过制作这些无线的程序在家中进行监测。这个机构的研究人员使用传感器来监测病人的脑部活动。如果传感器发现意外的脑部活动，它们给 10m 外的射频询问器发送一个警报。这个原型已经在比利时的 Universitaire Ziekenhuizen Leuven（UZ）进行了测试。

比利时的根特大学医院使用一个基于 RFID 真实时间的定位系统（RTLS）来为护士和其他医疗人员提供出现紧急情况的病人地点。这种集成 RFID 传感器网络在病人心脏病出现时监测并发送给医疗人员含有病人地点的警报。这种原型用到了航空侦察 T2 有源 WiFi 标签，能够把标签特有的 ID 传送给医院 WiFi 网络。

一种为家庭医疗监测和老年健康监测的集成 RFID 和 WSN 方案由 Ho 等[30]提出。这个系统的目的是监测老年人需要的药量，帮助他们服用精确的用量，作为 Intel 实验室对医疗原型的扩展。这个系统由一个高频 RFID 读卡器、一个超高频 RFID 读卡器、重量刻度、基站和三种模式组成。高频 RFID 标签放置在每一个药瓶上来识别每个瓶子，同时高频 RFID 读卡器用来监测它范围内的药瓶。一个药瓶的移动或者替换将会被定期地读取到相应的读卡器里。通过使用重量刻度与 RFID

嵌入在药瓶里的标签相结合，可以测定病人食用的药品以及药量。

每个病人也有一个 RFID 标签，病人通过相关联的 RFID 读卡器来提醒选择需要的药（通过声音或光警告）。这种原型是那种方法的延伸，已经被 Intel Labs[29] 提出，在此高频 RFID 读卡器和一些标签和两个传感器节点用来监测病人的药品摄入量。

国际西雅图研究中心和华盛顿大学也提出了一种智能家庭模型系统，叫做 “Caregiver's Assistant and CareNet Display”[22,67]，它能够发现、监测并记录老年人的日常生活活动，这是通过收集附着在家庭物品上如邮票大小无线标签的数据实现的。物品有关的信息和它们被触及的时间通过 WSN 被收集和传送。统计方法用在这个数据上来发现高水平的活动，填满来自国家授权的日常生活形式（ADL）的入口。这种原型系统的目的是帮助老年人来控制他们每天的活动，不需要一个二十四小时医疗人员。因此，医疗人员可以把焦点集中在老年人的医疗质量上而不是完成单调的工作。

Kim 等人[41]已经在医院的血液监测管理系统中部署了集成传感器标签。这个系统可以用来持续的监测血液冰箱库的温度以及跟踪血液包的位置。而且，这种系统能够用来阻止病人血液不搭配以及对血液包进行温度监测。这是通过使用血液库和贴在血袋上 RFID 传感器标签的传感器网络来实现的。Crossbow 技术 MTS420CA 传感器用来记录冰箱、血库和血袋的温度。TempSens KSWRFID 传感器标签用来标记血袋，它有足够的内存来存储 64 个温度测量数据。。

一种无源无线的 RFID 传感器能够植入到病人的食道壁上来发现胃逆流阻抗的变化是由 Ativanichayaphong 等[5]提出来的。集成传感器标签已经用来监测病人的牙齿健康。与一所联合国牙科学校合作的更精确仪器[31]在假牙上贴有 pH 传感器标签来监测病人嘴里食物的酸碱度[16]。

18.5.2 供应链管理中集成 RFID 和传感器网络

RFID 系统和传感器网络广泛应用于对存货清单，产品跟踪和资产监测的供应链，同时传感器网络也用在空间环境监测。然而，集成传感器网络的 RFID 系统打开了新的方向。RFID 系统能够精确地识别对象，但经常提供有关对象位置的不可靠信息。另一方面，传感器在确认对象位置方面有许多优点，但它们不能够对它进行成功的识别。RFID 和 WSN 的高效集成为精确位置跟踪提供了很大的优势。而且，集成 RFID 和 WSN 可以对产品以及危险环境条件进行监测，例如，易受损的或贵重的产品经不起的高温或者湿度。此外，集成 RFID 和 WSN 技术可以以一种便捷、便宜且无误的方式，在不用进行直接手工检查的情况下，（远距离）监测到环境和温度的自动变化。在供应链管理方面有一大批的可能集成 RFID 和 WSN 的方法。我们列举其中最重要的一部分。

一种集成 RFID 和 WSN 的自动资产追踪的方案已经由 McKelvin 等人[47]提出。这种方案集成无线传感器节点和 RFID 读卡器。更加精确的是，一个无线传感器节点被连接到主机（例如 ordinary PC），此处带标签产品的详细目录保存在数据库中。

另一种无线传感器节点集成在RFID读卡器中（读卡器节点）。主机节点的用户能够解决数据库中的查询，然后通过传感器网络转发给读卡器节点。这个问题接着传送给RFID读卡器，需要的数据被恢复。通信是双向的。因此，数据可以通过相同的接口从读卡器发送到主机设备。

PROMISE[53]是一个欧洲的方案旨在使用智能嵌入式设备监测和跟踪一个产品的生命周期。其中目标之一是使用集成的RFID标签和传感器检索信息并且随后提高产品制造、使用和回收的方法。此外，一些诸如InfraTab[35]和CliniSense[19]的公司集中于开发集成传感器标签来监测易腐烂食品的新鲜情况。它们关注于结合温度和持续时间来估计细菌增长的可能性。

18.5.3 其他应用

除了在医疗健康和供应链管理领域的应用，还存在许多其他现实世界集成RFID和WSN的方案。在本章中，我们将展示在科学和工程方面的应用方案，包括火灾监测、监测船舶载重、战争中武器状况以及管理牲口。在所有的方案中，集成RFID和传感器网络简化了程序，提高了效率。然而，所有方案中，不得不克服存在的不同技术挑战来实现集成系统顺利的运作。

集成RFID和传感器网络的一种应用是发生火灾时通知消防队员，这是由一个叫做Telexpath的墨尔本无线通信公司提出来的。这种方案以集成RFID芯片为基础，工作在433MHz，使用专有的空气接口协议和无线热量传感器。假如一个热量传感器感应了预先设定的2℃以内的温度，它将发送唯一的ID号给询问器。下一步是标签ID号的反向检查并且把这个通知发送到个人手机。因此，消防人员很迅速的注意到并且能够既快又高效的对火灾做出反应。这种方法也能够应用于警惕人们，不仅应用于火灾中也应用于其他危险火灾或者设备失败的情况下。

一种用来早期火灾监测的类似方法是由伯克利[23]美国加州大学的研究人员提出来的。这种方法以称作“Firebug”的使能（激活）GPS的无线传感器的发展为基础，它从传感器收集有关接近火灾的即时数据并通过包括Chipcon尘埃的射频标签传送这些数据。因此，消防人员能够获得火灾的速度和强度相关信息并作出恰当的反应。

Siemens IT Solutions and Services[73]控制着一个有关使用集成RFID和传感器网络技术的研究来监测船舶运输。它们已经证明实现从它们离开到抵达目的地实现对船舶进行持续的监测是可能的。这种方案使用有源RFID转发器和传感器。通过RFID收集的数据通过GSM或者FPRS通信网络传送给卫星通信设备。传送的信息可能包含船只的温度和位置。持续监测船只的优点是为有关订单的安排提供了重要的好处。例如，假如一个警报被触发，而商品仍然在海上，顾客可以注意到并执行新的订单。

在大不列颠的一个石油化学产品工厂里，无线寻呼机已经采用了集成RFID和WSN来高效的管理化学清单和提高存货的清晰度。这个方案涉及传感器标签和互相分享的数据的通信与协作。在这种集成RFID传感器网络中，每个RFID传感器

标签收集数据并把它们传送给网络中任何其他节点。使用的 RFID 传感器标签用来监测周围的环境，根据设定的规则提供警报。每个标签采用一个移动传感器，一个无线收发器，它使用专有的一对一协议进行互相通信。每个传感器标签除了传送自己的 ID 外，还有所有距离节点 3m 以内的环境状况的详细信息。

一种进行多重跟踪的混合的异构类型 RFID 系统和传感器结构由 Mori 等人[50]提出。这种系统集成分布式地面压力传感器和一个完整的 RFID 系统。通过使用地面压力传感器或者焦点热传感器，可以发现是否有人在某个地区。RFID 用来对人进行识别，因为当他们存在的地方重叠时，传感器就不能跟踪每一个人。实验结果证明这种方法可以在日常生活中对多个人进行高效地跟踪。

集成 RFID 传感器网络已经被 U. S. Navy[65]用来监测飞机重要部分的储藏条件。The U. S. Navy 和亚特兰大 Georae 技术机构合作采用一种 RFID 系统，它不需要 RFID 读卡器扫描每个标签，但是使用电池供电传感器标签，能够使他们相互通信和传送信息。数据从终端转发器传送给单独的读卡器。RFID 传感器标签能够测量温度、湿度和空气压强来与其他 RFID 传感器标签通信。RFID 传感器标签是自己供能，有两年的寿命。标签只有在被基站询问后才传送信息，这同样需要发送一个详细的安全代码。因此，将会阻止可能泄露的信息给未经授权的敌人，这些信息会揭露有关船只位置的敏感信息或者其他敏感的存货清单信息。

HP 也在叫做智能 LOCUS 和智能 Rack[51]的两种原型应用中使用集成的 RFID 传感器网络。智能 LOCUS 控制和监测一个含有录像机和读卡器的传感器网络。覆盖在网络上的传感器把 RFID 读卡器和便宜的录像机通过 802. 11b 网络连接起来。例如，摄像机提供仓库中物品移动的相关信息。

在智能 RACK 方案中，热量传感器和高频 RFID 读卡器用来监测服务器柜子的温度。在行李架中每个架子上装备一个 RFID 读卡器，用专门的芯片存储机器的唯一 ID 号读取来自服务器的高频信号。RFID 读卡器使用于每个服务器上，并且 14 根天线配置在服务器门上来读取贴在每个服务器柜子里 13. 56MHz 的 RFID 标签。五六个热传感器通过电线连接到柜子门上并与读卡器相连。热传感器监察位于服务器柜子的服务器的温度。高频 RFID 读卡器和传感器是网状的，收集来的数据用来即时产生 2in 图像，它显示的是每个柜子轮廓的温度。这个应用对于使用大量服务器的公司来说是必要的。一旦出现不正常的高温，将会发送一个警报并且通知人们可能出现的问题。

另一个集成 RFID 和 WSN 的应用是在战场上[36]。最近，一种可以嵌入到武器里跟踪在战斗中战士使用武器射击次数的原型已经生产出来了。因此，军队可以估计何时每个武器达到了它的极限。这个原型集成传感器于 RFID 标签中。这是通过在每个武器中放置压电传感器，微型处理器和 RFID 标签实现的。这个压电传感器通过感知后冲力能够决定什么时候射击了武器。微型控制器的内存有限，能够记录和存储传感器的输出。最终，RFID 标签用来把数据传输给 RFID 读卡器。这个原

型能够跟踪射击的数量和次数，也能够推测出开火的强度，诸如加速、激烈和电磁共振频率。所有这些特性能够帮助来估计武器的有效时间。

集成 RFID 和 WSN 已经用来感应坦克的油量[8]，来监测用在原油精炼中的大型制冷风扇的震动情况，也用来在形成汽油期间监测和减轻混合不同原油水平的难度。所有的三种应用是试验性的方案，由使用 i-Sense Talker 设备的 Sense-Comm 来执行，此设备是 iRFID 标签的先进版本，包含的传感器不同于那些通常使用在 iRFID 标签的传感器。集成 RFID 传感器标签将很大程度上使上面提到的原油精炼过程更加方便。例如，用传统的方法来监测坦克中液体的含量使用一个机械的测量仪表，它需要人工的检查或者浸泡一个有线的传感器于坦克中。然而，考虑到这些区域都是危险的并且不管是用有线还是人工控制执行这些工程都是很昂贵的，所以这两种都不是最佳的方案。

另一种集成 RFID 和 WSN 的应用是对牲口的便捷管理。Zigbeef[70] 是市场硬件，能够被大农场主用来更容易的管理他们的牧群。Zigbeef 使用集成 RFID 传感器标签，能够感知动物的移动并把这些数据传送给 RFID 读卡器。这些数据也能够用在骑术表演中，传感器可以贴在表演的牛身上。因此，使用移动读卡器，观众将能够观察到有关在公牛把骑士扔出去之前跳起的程度的数据。ZigBeef 也正在计划用网络能力扩展当前的系统，使 ZigBeef 标签将能够从一个向另一个传递数据。因此，读的范围将会根据有多少的牲口穿梭在这个场所来扩展，最终，数据将被发到读卡器。

另一个集成 RFID 传感器网络的应用是以 RFID 为基础的内存助手的发展，它将会在人离开屋子而没有带必备的物品时告诉人们。所有必需的物品都被标记了，同时一个压力传感器发现她存在屋子的前门上，如果有些东西丢失了就会激活 RFID 读卡器。另一个方案包括使用温度传感器贴在母牛的胃里[60]来预测小牛的出生。这个系统以测量母牛的身体温度为基础，可以在它的牛犊出生之前 24h 监测母牛的体温以免发生意外。

日本国内事务交流部也发明了一个系统，以集成传感器标签为基础来收集灾难地区的信息[61]。传感器标签从直升机上洒下来，从灾难地区收集重要的信息，包括灾难受害者中可能幸存的人。

一种为群体旅游系统的集成 RFID 和 WSN 的方案由 Chen 等人[12]提出。根据这种集成方案，每一个群体的导游拿着一个标记，它能够发射 4kHz 的信号。每个组的成员携带着附有无源 RFID 标签的票，它存储在群体的 ID 中。传感器节点贴在方向板上来显示简单的引导方向。导游标记发出的信号被无线传感器节点所识别，因此导游的位置可以被追踪。传感器网络中的一些节点选作为帮助中心，连接着笔记本和 RFID 读卡器。

集成 RFID 和 WSN 也是自动 ID 实验室宇宙航空航天的研究方向。通过在航天器上使用该系统，可以准确地获取航天器的配置。通过预测，可以方便地对航天器进行维护和修理。邮寄顾客相关的设备诸如订票、加工行李和食物链也被认为是使

用集成 RFID 和 WSN 的研究方向。

18.6 结论和开放性问题

毫无疑问，集成 RFID 和 WSN 即将发生，它将产生高水平的合作和更多的技术改进。这种集成带给我们的好处不仅是揭示对象的位置和身份，也有它当前的状态。这些集成网络将扩展传统的 RFID 系统，在环境控制和工业生产中给我们带来重要的好处。然而，必须要付出更多的努力来实现高效的集成 RFID 和 WSN。

广泛采用集成 RFID 和 WSN，对于解决一些公开的问题和挑战是很重要的，例如在大型集成 RFID 网络和 WSN 中减少可能的干扰，因为在网络中无线设备的数量越多，潜在的干扰可能就越大。因此，为 WSN 和 RFID 定义一个防冲突的清单是很重要的。

对于广泛使用集成 RFID 和 WSN 的一个重要步骤是部署工具、方法和方式，它们能普遍的使用在大部分应用中。然而，当配置这些工具方法和标准时很重要的是考虑有效资源的限制。此外，先进的方法应该相当普遍，使他们进化成为标准是可行的。

这个方向令人感兴趣的首创精神是 IntelliSense RFID 方案[72]，它是北欧研究程序 NORDITE 的一部分。IntelliSense 的目标是发展 RFID 设备的多重协议，它能够与不同的通信协议工作在不同的频率带宽下。因此，将会发明二重带宽传感器标签，并且能够用在各种各样的设备中。例如，多重带宽传感器标签可能被超高频 RFID 读卡器用在后勤应用，顾客通过集成于手机的高频 RFID 读卡器来检索存储在内存中的数据。

除此之外，集成 RFID 和 WSN 技术的模拟装置在广泛的部署与集成 RFID 和 WSN 的研究是很重要的。此外，考虑到市场受到成本的驱使，成功的集成 RFID 和 WSN 将很大程度上依靠最低的原料成本和简单最有效的制造过程。

参 考 文 献

1. Adage Solutions, Products, http://www.adage.se/Adage/Page____240.aspx, Accessed July 2009.
2. Aeroscout Enterprise Visibility Solutions, AeroScout T3 tags, http://www.rfidglobal.org/product/2007_7/aeroscout_t3_tags.html, Accessed July 2009.
3. Alien Technology, Alien Technology ALB-2484 tag, http://www.rfidsolutionsonline.com/ecommcenters/alien.html, Accessed July 2009.
4. Alr-9770 Series Multi-Protocol RFID Readers, 2005, http://www.directouchpos.com/products/downloads/alr9774.pdf, Accessed July 2009.
5. T. Ativanichayaphong, J. Wang, W.-D. Huang, S. Rao, H.F. Tibbals, S.-J. Tang, S.J. Spechler, H. Stephanou, and J.-C. Chiao, Development of an implanted RFID impedance sensor for detecting gastroesophageal reflux, in *Proceedings of IEEE International Conference on RFID*, Grapevine, TX, pp. 127–133, March 26–28, 2007.
6. B. Bacheldor, Belgium hospital combines RFID, sensor to monitor heart patients, RFID Journal, March 2007, http://www.rfidjournal.com/article/articleview/3120/1/1, Accessed July 2009.

7. B. Bacheldor, Fighting fires with RFID and wireless sensors, *RFID Journal,* November 2006, http://www.rfidjournal.com/article/articleview/2799/1/1, Accessed July 2009.
8. B. Bacheldor, Oil refineries to test sensor tags, *RFID Journal,* January 2007, http://www.rfidjournal.com/article/articleview/3006/1/1, Accessed July 2009.
9. Life Chip. Bio-Thermo, http://lifechip.com.au/products.php?id=4, Accessed July 2009.
10. Bisa Technologies, 2.4 GHz temperature sensor tag (24TAG02T), http://bisatech.com/product.asp?pid=2&zid=3&do=view&id=49, Accessed July 2009.
11. Bisa Technologies, 2.4 GHz vibration sensor tag (24TAG02V), http://bisatech.com/product.asp?pid=2&zid=3&do=view&id=50, Accessed July 2009.
12. P.Y. Chen, W.T. Chen, C.H. Wu, Y.C. Tseng, and C.F. Huang, A group tour guide system with RFIDs and wireless sensor networks, in *Proceedings of the 6th International Conference on Information Processing in Sensor Networks,* Cambridge, MA, pp. 561–562, April 25–27, 2007.
13. J. Cho, Y. Shim, T. Kwon, and Y. Choi, SARIF: A novel framework for integrating wireless sensor and RFID networks, *IEEE Wireless Communications,* 14(6), 50–56, December 2007.
14. N. Cho, S.-J. Song, S. Kim, S. Kimn, and H.-J. Yoo, A 5.1– W UHF RFID tag chip integrated with sensors for wireless environmental monitoring, in *Proceedings of the 31st IEEE European Solid-State Circuits Conference (ESSCIRC'05),* Grenoble, France, pp. 279–282, 2005.
15. R. Clauberg, RFID and sensor networks, in *Proceedings of RFID Workshop,* University of St. Gallen, Switzerland, September 2004.
16. J. Collins, Passive tag powers sensor networks, switches, *RFID Journal,* April 2005, http://www.rfidjournal.com/article/view/1520/1/1, Accessed July 2009.
17. J. Collins, BP tests RFID sensor network at U.K. plant, *RFID Journal,* June 2006, http://www.rfidjournal.com/article/view/2443/1/1, Accessed July 2009.
18. J. Collins, ShyeTek shrinks the RFID reader, *RFID Journal,* January 2004, http://www.rfidjournal.com/article/articleview/778/1/1, Accessed July 2009.
19. J. Collins, Sensing a product's shelf life, *RFID Journal,* April 2005, http://www.rfidjournal.com/article/view/1539/1/1, Accessed July 2009.
20. Crossbow, Mica2 wireless platform, http://www.xbow.com/Products/productdetails.aspx?sid=174, Accessed July 2009.
21. H. Deng, M. Varanasi, K. Swigger, O. Garcia, R. Ogan, and E. Kougianos, Design of Sensor-Embedded Radio Frequency Identification (SE-RFID) systems, in *Proceedings of IEEE International Conference on Mechatronics and automation,* Henan, China, pp. 792–796, June 2006.
22. E. Dishman, Inventing wellness systems for aging in place, *IEEE Computer Magazine,* 37(5), 34–41, May 2004.
23. D.M. Doolin and N. Sitar, Wireless sensors for wild re-monitoring, in *Proceedings of SPIE Symposium on Smart Structures & Materials NDE 2005,* San Diego, CA, March 6–10, 2005.
24. C. Englund and H. Wallin, RFID in wireless sensor networks, Master thesis, Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, Goteborg, Sweden, April 2004.

25. EPC Global, Class 1 Generation 2 UHF air interface protocol standard "Gen 2" http://www.epcglobalinc.org/standards/uhfc1g2/uhfc1g2_1_1_0-standard20071017.pdf, Accessed July 2009.
26. Evidencia, ThermAssureRF wireless temperature recorder, http://www.evidencia.biz/products/acti-tag.htm, Accessed July 2009.
27. R.B. Ferguson, Gentag patent adds RFID sensor network feature to mobile devices, December 2006, http://www. eweek. com / c /a/ Mobile-and- Wireless/Gentag-Patent-Adds-RFID-Sensor-Network-Feature-to-Mobile-Devices, Accessed July 2009.
28. A. Ferrer-Vidal, A. Rida, S. Basat, L. Yang, and M.M. Tenzeris, Integration of sensors and RFID's on ultra-low-cost paper-based substrates for wireless sensor networks applications, in *Proceedings of the 2nd IEEE Workshop on Wireless Mesh Networks (WiMesh 2006)*, Reston, VA, pp. 126–128, September 25, 2006.
29. K. Fishky and M. Wang, A flexible, low-overhead ubiquitous system for medication monitoring, Intel Research Technical Report IRS-TR-03-011, October 2003.
30. L. Ho, M. Moh, Z. Walker, T. Hamada, and C.-F. Su, A prototype on RFID and sensor networks for elder healthcare: Progress report, in *Proceedings of the 2005 ACM SIGCOMM Workshop on Experimental approaches to Wireless Network Design and Analysis, Applications, Technologies, Architectures, and Protocols for Computer Communication*, Philadelphia, PA, pp. 70–73, 2005.
31. Instrumentel Telemetric Technologies, Tag sensor: ICT tag sensor, http://www.instrumentel.com/specs/Tag%20INTRA%20SENSE.pdf, Accessed July 2009.
32. International Organization for Standardization (ISO), ISO 11785: 1996—Radio frequency identification of animals, http://www.iso.org/iso/catalogue_detail?csnumber=19982, Accessed July 2009.
33. International Organization for Standardization (ISO), ISO 15693-2: 2006, Identification cards—Contactless integrated circuit cards—Vicinity cards—Part 2: Air interface and initialization, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39695, Accessed July 2009.
34. C. Ilic, Using tags to make teeth, *RFID Journal*, October 2004, http://www.rfidjournal.com/article/view/1206/1/1, Accessed July 2009.
35. InfraTab, Monitor, track and trace perishables, http://www.infratab.com, Accessed July 2009.
36. J. Jackson, Ready, aim, record: Army's prototype system uses RFID tags to track weapons use, *GCN government Computer News*, May 2008, http://www.gcn.com/Articles/2008/05/01/Ready-aim-record.aspx, Accessed July 2009.
37. Machine Talker, iRFID—Intelligent radio frequency identification device, *Q3 2006, Preliminary Specification*, http://www.machinetalker.com/pdf/iRFID-Q3-2006-Spec.pdf, Accessed July 2009.
38. Microsensys, Telid 210 passive sensor TAGs, http://www.microsensys.de/products/sensors/TELID210/TELID210.html, Accessed July 2009.
39. Microsensys, Telid 310 passive sensor TAGs, http://www.microsensys.de/products/sensors/TELID310/TELID310.html, Accessed July 2009.
40. S. Kim, J.-H. Cho, H.-S. Kim, H. Kim, H.-B. Kang, and S.-K. Hong, An EPC Gen 2 compatible passive/semi-active UHF RFID transponder with embedded feram and temperature sensor, in *Proceedings of IEEE Asian Solid-State Circuits Conference*

(ASSCC'07), Jeju, Korea, pp. 135–138, November 12–14, 2007.
41. S. J. Kim, S. K. Yoo, H. O. Kim, H. S. Bae, J. J. Park, K. J. Seo, and B. C. Chang, Smart blood bag management system in a hospital environment, *Personal Wireless Communications*, Springer, Berlin/Heidelberg, Vol. 4217/2006, pp. 506–517, September 2006.
42. D. Kiritsis, Ubiquitous product life-cycle management using product embedded information services, in *Proceedings of International Conference in Intelligent Maintenance Systems (IMS'2004)*, Arles, France, July 2004.
43. H. Kitayoshi and K. Sawaya, Long range passive RFID-tag for sensor networks, in *Proceedings of 62nd IEEE Vehicular Technology Conference (VTC'05)*, Dallas, TX, Vol. 4, pp. 2696–2706, September 25–28, 2005.
44. KSW Microtec, Active RFID—VarioSens, http://www.ksw-microtec.de/index.php? ILNK=Active_RFID_VarioSens&iL=2.
45. H. Liu, M. Bolic, A. Nayak, and I. Stojmenovie, Integration of RFID and wireless sensor networks, in *Proceedings of Sense IP 2007 Workshop at can SenSys*, Sydney, Australia, November 6–9, 2007.
46. A. Mason, A. Shaw, A.I. Al-Shamma'a, and T. Welsby, RFID and wireless sensor integration for intelligent tracking systems, in *Proceedings of 2nd GERI Annual Research Symposium GARS-2006*, Liverpool, U.K., June 2006.
47. M. L. McKelvin, M. L. Williams, and N. B. Berry, Integrated radio frequency identification and wireless sensor network architecture for automated inventory management and tracking applications, in *Proceedings of the 2005 Conference on Diversity in Computing*, Albuquerque, NM, pp. 44–47, 2005.
48. Mica2dot Series, http://www.willow.co.uk/html/mpr5x0-_mica2dot_series.html, Accessed July 2009.
49. J. Mitsugi, T. Inaba, B. Patkai, L. Theodorou, J. Sung, T. Sanchez Lopez, D. Kim, D. McFarlane, H. Hada, Y. Kawakita, K. Osaka, and O. Nakamura, Architecture development for sensor integration in the EPCGlobal network, *White Paper WP-SWNET-018, Auto-ID Labs*, July 2007.
50. T. Mori, Y. Suemasu, H. Noguchi, and T. Sato, Multiple people tracking by integrating distributed floor pressure sensors and RFID system, in *Proceedings of 2004 IEEE International Conference on Systems, Man and Cybernetics*, Vol. 6, pp. 5271–5278, The Hague, the Netherlands, October 10–13, 2004.
51. M. C. O'Connor, HP kicks off us RFID demo center, *RFID Journal*, October 2004, http://www.rfidjournal.com/article/articleview/1211/1/50, Accessed July 2009.
52. School of engineering and Harvard University Applied Sciences, Wireless sensor networks for medical care, September 2006, http://fiji.eecs.harvard.edu/CodeBlue, Accessed July 2009.
53. PABADIS, PABADIS based product oriented manufacturing systems for re-configurable enterprises, http://www.uni-magdeburg.de/iaf/cvs/pabadispromise/beschreibung_english.pdf.
54. B. Patkai and D. McFarlane, RFID-based sensor integration in aerospace Technical Report AEROID-CAM-009, Auto-ID Lab, University of Cambridge, U.K., November 2006.
55. Patent storm, U.S. Patent 7125382 – Embedded Bio-Sensor system, October 2006, http://www.patentstorm.us/patents/7125382/fulltext.html, Accessed July 2009.

56. Phase IV Engineering Inc., SensIc RFID ASIC, http://www.phaseivengr.com/Application_PDFs/61-100005-00_Rev_1_0_SensIC_RFID_ASIC.pdf, Accessed July 2009.
57. A.G. Ruzzelli, R. Jurdak, and G.M.P. O'Hare, On the RFID wake-up impulse for multi-hop sensor networks, in *Proceedings of 1st ACM Workshop on Convergence of RFID and Wireless Sensor Networks and their Applications (SenseID) at the Fifth ACM Conference on Embedded Networked Sensor Systems (ACM SenSys 2007)*, Sydney, Australia, November, 2007.
58. RFID anywhere overview, 2005, http://www.sybase.com/content/1034553/rfidanywhereoverview.pdf.
59. RFID in Japan, RFID-based alert when you leave home, March 2006, http://rfidinjapan.wordpress.com/2006/03/24/rfid-based-alert-when-you-leave-home/, Accessed July 2009.
60. RFID in Japan, RFID tags in cow's stomach predict child birth, March 2006, http://rfidinjapan.wordpress.com/2006/03/23/rfid-tags-in-cows-stomach-predict-child-birth/, Accessed July 2009.
61. RFID in Japan, Sprinkling RFID sensor tags from the sky, March 2006, http://rfidinjapan.wordpress.com/2006/03/20/sprinkling-rfid-sensor-tags-from-the-sky/, Accessed July 2009.
62. RFID in Japan, RFID system for repelling monkeys, June 2006, http://rfidinjapan.wordpress.com/2006/06/27/rfid-system-forrepelling-monkeys/, Accessed July 2009.
63. RFID in Japan, Battery-less sensor tags, http://rfidinjapan.wordpress.com/2006/04/20/battery-less-sensor-tags/, Accessed July 2009.
64. A. Rida, R. Vyas, S. Basat, A. Ferrer-Vidal, L. Yang, S.K. Bhattacharya, and M.M. Tentzeris, Paper-based ultra-low-cost integrated RFID tags for sensing and tracking applications, in *Proceedings of the 57th Electronic Components and Technology Conference (ECTC'07)*, Reno, NV, pp. 1977–1980, 2007.
65. M. Roberti, Navy revs up RFID sensors, *RFID Journal*, June 2004, http://www.rfidjournal.com/article/articleview/990/1, Accessed July 2009.
66. A.P. Sample, D.J. Yeager, P.S. Powledge, and J.R. Smith, Design of a passively-powered, programmable sensing platform for UHF RFID systems, in *Proceedings of 2007 IEEE International Conference on RFID*, Gaylord Texan Resort, pp. 149–156, March 26–28, 2007.
67. M. Philipose, S. Consolvo, T. Choudhurg, K. Fishkin, M. Perkowitz, I. Smith, D. Fox, H. Kautz, and D. Paterson, *Demonstration in the Sixth International Conference on Ubiquitous Computing (UbiComp'04)*, 7–10 September 2004, Nottingham, England.
68. Sensitech, ColdStream infrastructure: Integrated RF-enabled temperature monitoring infrastructure, Sensitech Cold Chain Visibility, http://www.sensitech.com/PDFs/applications/ColdStream_InfraStructure.pdf, Accessed July 2009.
69. SkyeTek, Inc., SkyeTek RFID Readers SkyeRead M1-Min, Product Reference Guide, 2004, http://www.ece.osu.edu/~cglee/ECE682Y/Doc/SkyeModule_M1-mini_Reference_Guide.pdf, Accessed July 2009.
70. C. Swedberg, ZigBeef offers ranchers a long-distance cattle head count, *RFID Journal*, http://www.rfidjournal.com/article/articleview/3935/1/1, Accessed July 2009.

71. C. Swedberg, Dutch researchers focus on RFID-based sensors for monitoring apnea, epilepsy, *RFID Journal*, December 2007, http://www.rfidjournal.com/article/articleview/3780/1/1, Accessed July 2009.
72. O. Vermesan, N. Pesonen, C. Rusu, A. Oja, P. Enoksson, and H. Rustad, IntelliSense RFID—An RFID platform for ambient intelligence with sensors integration capabilities, *ERCIM News*, 67, 40–41, October 2006.
73. R. Wessel, Cargo-tracking system combines RFID, sensors, GSM and satellite, *RFID Journal*, January 2008, http://www.rfidjournal.com/article/articleview/3870, Accessed July 2009.
74. G. Yang, M. Xiao, and C. Chen, A simple energy-balancing method in RFID sensor networks, in *Proceedings of 2007 IEEE International Workshop on Anti-Counterfeiting, Security, Identification*, Xiamen, China, pp. 306–310, April 16–18, 2007.
75. L. Zhang and Z. Wang, Integration of RFID into wireless sensor networks: Architectures, Opportunities, in *Proceedings of the Fifth International Conference on Grid and Cooperative Computing Workshops (GCCW'06)*, Changsha, China, pp. 463–469 October 2006.
76. S.-H. Zhou and N.-J. Wu, UHF RFID Front-end with magnetic sensor, in *Proceedings of the 8th International Conference on Solid-State and Integrated Circuit Technology*, Shanghai, China, pp. 554–556, October 2006.

第19章 应用于智能家居系统的RFID与无线传感器网络的集成

科技的快速发展在短时间内给现实生活带来了许多变化。因此，智能家居的梦想将在不久的将来变得更加可行。诸如无线传感器网络和无线射频识别技术有着引人注目的特点，使它们成为环境中的伟大的候选人，并且很大程度上受益于彼此。“智能家居”包括各种现在以及将来生活和工作方面的各种方法。其目标涵盖的范围从增加日常生活的舒适性到为老年人和残障人士提供更加独立的生活。“普适计算”由 Mark Weiser 在他的随笔《The Computer for the 21st Century》中提出，描述的是计算机和信息技术的普遍性。注入日常生活的“智能物体”的任务是使用各种传感器感知即时环境并处理这些信息。此功能把一种人工智能分配给普通的、众所周知的对象，能够对信息进行全面的处理以及与几乎所有日常物品进行相互连接。可见和隐藏的信息从“穿戴计算机”和“智能服”到“智能”人工替代物变化。它通过扩展认知能力和信息处理能力在几乎所有部分支持用户，试图补偿某种缺陷。有关智能家居的挑战，尤其是赡养老人和残障人士，是弥补他们的缺陷，尽可能给他们更加独立的生活。在这一章节中，将会提出一种普遍的智能家居环境结构，描绘并最终评估一种实验性的装置。这种结构主要包括一个传感器网络和无源 RFID 标签，前者通过移动机器人来加强起作用，后者反过来补充传感器网络的功能。

19.1 概述

静态无线多跳网络（SANET）由传感器和执行器节点组成（通常通过无线技术通信），执行分布式感知和驱动任务。最近几年，对 SANET 方面的需求和兴趣正在增加。显而易见，SANET 是混杂的网络，拥有不同的传感器和执行器节点特性[4,5]。基本概念是传感器节点提供一个通信的基础结构，收集和发布有关物理环境的信息[12]。它们也可以在某种方式下用于定位和航海问题，部分在[5,24]中描述；然而，这不是本章的焦点。执行器节点作出的决定是以从发布节点或者本地系统获得信息为基础的。例如，它们可以给传感器网络提供管理和维护服务，诸如修补任务或者能量恢复。除此之外，执行器节点还需要提供可移动性，例如，通过可移动传感器控制移动性来改进网络。我们认为在许多情况下，更多的能量执行器节点诸如移动机器人需要用来满足更复杂的任务。这些机器人是自动的机器，可以在特定的环境中移动[6,20]。一种普遍的机器人是轮式机器人。它们可以完成执行各种任务

来改进和支持传感器网络。我们认为，维护和优化配置移动机器人对传感器网络的帮助如维护和优化部署传感器网络对机器人的帮助如提供位置信息，必须加以区别。

依照惯例，项目经常用一个用来识别项目的条形码来标注。替代条形码的一种选择是 RFID 标签。基本上，一个 RFID 标签代表的是能量有限、有计算能力和无线通信能力的发射机的简单芯片。然而，通常情况下，有源和无源 RFID 标签需要区分开来，我们专注的无源标签不需要一个专门的电池。相反，一个外部能源通过有限但足够发射信息的能量来无线支持 RFID 标签。信息中包含一个唯一的叫做电子产品编码（EPC）的数字，并存储在小型计算机的外部 RFID 读卡器的 EEPROM 中。一些智能家居环境中 RFID 的应用和相关安全威胁在参考文献［14］中提出来。RFID 能够用来识别产品、地点、宠物甚至人。RFID 标签的一个主要优点是读的过程不需要瞄准线位置，而条形码需要瞄准线排成直线。唯一的要求是 RFID 标签必须在距离读卡器的某个范围内。这个范围很大程度上取决于读卡器的天线。使用 RFID 技术，上百个 RFID 标签能被同时读取。RFID 技术正在使用的一个最大领域是提供连锁管理[30]。

RFID 的其他应用方案也专注于智能家居的应用，包括追踪频率丢失对象（FLO），诸如密钥，活动监测（例如识别人），也包括对 SANET 的支持（例如定位和不受限制的信息传输）。如果所有这些应用能够使用 RFID，将变得更加简单。在智能家居的场景下，已经显示 RFID 技术提供了许多利益。例如，基于 RFID 的自感知场所已经使用所谓的智能插座分析过。同样，针对家居中使用简单而普遍的传感器的活动识别已经通过调查来支持更复杂的智能家居场景[28]。Trumler 等人发明了一个智能门牌号，使用 RFID 技术作为街区的普遍环境[29]。最后，一个可编程的普遍空间已经被提出来，基于混合的有源和无源 RFID[23]。这些例子清楚地概括了 RFID 支持普遍环境的益处。在所有这些场景中，一个中心计算机连接和组织本地网络是很重要的，这一章中提出的场景中，这个中心计算机包含 RFID 读卡器、无线传感器节点和移动执行器诸如机器人系统。

“智能家居”一词覆盖了各种各样实践和理论的方法，处理今天以及将来的生活、居住和工作[32]。任务之一是实现自动化和解决日常生活中不同领域的问题，诸如家庭娱乐和健康医疗，并使这些单个的解决方法结合成为一个全面的网络整体。另一个任务是在全局任务范围内联合不同场景中各种单一的解决方案。所有显而易见的混杂的网络和部署的设备应该与巨型网络相互连接，覆盖整个居住区域，反过来连接到外部的全球性网络，例如因特网。整体的任务是收集并分析数据以一个自我组织自我控制的方式回应。所以，基本不需要用户的交互行为，日常任务就可以解决，在某种方式下环境或者家庭中似乎是“智能”的。20 世纪 90 年代初，各种概念和标准像 Konnex（KNX 技术）、欧洲安装总线（EIB）、欧洲的主页系统（EHS）以及其他许多都是通过不同的方式演化的。但是这些概念是基于用户交互

的。这种交互以一系列规则呈现；它们各自指定一种方法来回应。对于控制插座、灯光和其他内部或者外部建筑的系统都是可能的，但是没有内在的智能。随着科技的不断进步，使用更多人工智能概念替代控制机械装置的可能性在增加。过去几年，提出了各种方法，它们各自提供了一个单一的解决方法。所以，无日期可以命名为第一个和原始智能家居。一个相关的众所周知不同方面的项目一个小的选择是：

- 杜伊斯堡 InHaus（弗劳恩霍夫研究所）[27]
- LIVEfutura（弗劳恩霍夫研究所）[22]
- T - Com - Haus（柏林的 T - Com）[3]
- Futurelife（Huenenberg/Zug，Beisheim Group Metro）[8]
- 慕尼黑的各种项目（BW 大学/慕尼黑/微软）[7,21,25]
- OnStar 家居（底特律，国际互联网联盟）[11]
- TRON 的智能住宅（Nishi Azabu，K. Sakamura）[26]
- 智能医学家庭研究实验室（美国罗切斯特大学）[1]
- SENTHA（柏林数家德国大学和研究组）[19]
- INGA（创新网络建设自动化）[2]

在下一步中，研究将走向人工智能、自学习和自组织系统。智能对象呈现的环境不再受中心控制管理。它们收集并分析数据的同时几乎自动工作。所以可以找出在我们生活的所有方面的（行为）模式，像消费、睡觉等。用户不会感觉到依赖技术，或是被观察和控制，这些都需要显而易见的技术和隐藏的复杂性，尤其针对许多不同的用户，例如老人/年轻人，健康人/残障人士，对技术改进感兴趣的人/不想或者不能够通晓那些细节的人。

另一个目标是迅速做出对每个人都负担得起的、有用的改进。因此，新技术不得不和现有生活理念相结合使它成为我们日常生活的常见部分。直到现在，没有设计能被认为是最好的或者最有革新精神的，因为它们都遵循不同的方法和各种不同的目标。最常见的场景是覆盖日常生活的舒适和困难、娱乐方式和家庭娱乐，以及在人变老或者残障需要药物支持时保持独立。第一类别目标或者经济目标是使生活更舒心，同时减少成本。就工作进程的提高，它是很有活动性的，这不在本章介绍。

这一章的目的是帮助老年人或者残障人士尽可能独立生活。这些人是智能家居环境相关的一个重要的目标群体，所以需要考虑一些特别的需求。例如，必须建立健康监测和紧急情况帮助系统。控制基础设施和接口的需求必须容易和自说明。用户应该被结合起来并且在新的环境中感觉很好。当为这个目标群体发展智能环境时，主要关注于补偿残障和限制。在工作中，我们在现实环境中建立并且测试 SA-NET 来完成具体任务，例如监测和控制本地系统、家庭娱乐以及健康医疗。根据收集的数据，诊断中心能够分析监测病人的行为。例如，如信号灯能够被识别、分析

和控制。对于医疗健康应用，收集的数据可能相关。对于传感器网络，操作易受某种技术限制。因此，传感器节点只是有条件的可靠。进一步地说，维修、管理和诊断的新的功能应该结合起来。所以，部署移动机器人似乎是一个很适合的方法。它们可以证实传感器节点、分派任务、再编程传感器节点测量的数据，并局部化频繁丢失的东西（FLO），例如密钥和一些其他合作性任务。除此之外，移动机器人能用作智能家居环境的基站。我们的工作包括建立测试环境展示智能家居中提到的功能，测试的环境是网络实验室。它的发展是以传感器节点（Btnodes）、各种型号的射频识别芯片和移动机器人（Robertino）为基础的。像路由、感应机器人通信和机器人控制的基础作用在之前的工作中部分已经发展了起来。

19.2　我们的家居智能环境

下面介绍一种智能家居的应用场景，假设一位老人拥有并居住在一个普通公寓。随着感知和记忆能力不再那么好，在许多方面可能受到阻碍。对于一个老年人或者残障人士，处理日常任务将会更加困难，像控制暖气和空调，或者甚至一个简单的家庭娱乐系统。年龄的不断增长是短时期内这个项目的重要迹象以及数据应该被不断地核对的一个原因。人也不能够在他需要或者想要的任何时间看内科医生，因为事实上他自己无法去看医生。因此，一种解决方法必须被提出来使病人和医生以一种舒适的方式交谈。医生有可能远程观察重要迹象和数值，与病人联系，就像正常的做法一样。除此之外，家庭成员可能在很远的地方，而附近经常没有看护。因此，以防紧急情况发生，一个可靠的紧急系统可以通知医生和救护车，并且立即把突发事件通知给家属。

19.2.1　目标

特别关注老年人或者残障人士，更高水平的目标是尽可能弥补他生活中的各种限制，来确保病人尽可能更加独立的生活。现在记住所有的任务，为一个智能家居场景列一个基本的目标是可能的。因此，指定的首要目标是感知和监测系统的发展，它接管部分看护对象的浪费感知，如果必要，维持他的记忆能力。详细地讲，构架必须照顾到家庭系统、空调、灯和暖气装置，也控制家庭娱乐和安全系统的基本作用。为了与医生相互联系，并获得医生实时检查的重要的监测迹象和数据，作为一个额外的目标，构架需要聚集数据，局部分析数据，使数据容易得到，开启紧急呼叫并且提供录像/声音通信。除此之外，移动机器人应该用来保存和管理网络相关的任务，为了移动传感器的全部任务分配和局部化FLO。

这些目标需要使用相当多的技术，但是，这种技术必须是清楚的并且容易使用，这个事实是最为重要的观点之一。如果它们没有被用户接受，那么所有的改进和技术是无用的。因此，用户系统界面必须是尽可能简单和强大，并且这种系统必

须基本上以自组织的方式运行。

19.2.2 现实需求和实验室限制

为了把已经提出的构架描绘成一个现实的图画，若干规定需要应用到部件和组件以及环境和用户中。这个计划要把给定的预算和可能性合并起来。所有的设备要提供必要的界面在整个网络中相互连接。即使它们的设计不得不支持它们预定的范围。在公寓和家里，规定不同于那些办公室和工厂建筑。小且令人满意的已建设备对于某种舒适是必要的。需要连接高速因特网来满足各种服务的需要，像声音/录像技术和交互式的医疗护理。就针对老年人和残障人士的智能家居而言，一个备份的因特网连接是保证紧急系统可靠性所必需的。即使网络垮掉了，人们应该能够紧急呼叫。随着（W）BAN 系统研究的快速发展，这种可能性也在提高，因此，多功能和更令人兴奋的技术在不久的将来以某种可见的方式是可以达到的，像传感器膏药或植物。直到现在，通常使用在（W）BAN 中的技术单元和设备有时在可见的所有时间会干预平常的生活。布置的要求也要根据环境和用户，就像公寓被认为不管是建立有线还是无线网络都是很合适的。为了部署移动机器人，地板也不得不被改变。用户必须愿意使用和理解某种低水平的技术去控制它，即使努力来尽可能简化过程和使用性。

为了在我们实验室演示一个场景，一些限制已经明确从而使用现有的硬件和软件作为一个基础来实施一个描绘了设计的通用构架的智能家居测试环境。显而易见的是，环境是受限的，或者在某些方面是相当简化的。与公寓或者家里相比，实验室配备有一些椅子、桌子和实验设备。除此之外，工作地点和其他技术器材也要考虑。因此，一般的居住空间的概念必须被抽象。在预算之内，建立一个拥有所有指定设备的环境是不可能的，如家庭环境，家庭娱乐，安全系统和空调。其他诸如暖气装置和灯系统由大学维修服务。专用服务器提供代理和 Web 服务。在有效地硬件内，任何执行禁令的组件包括在内。所使用的 RFID 读取器提供了最大 10cm 的读取距离，这意味着 RFID 标签应该放在地上。他们也应该以一种方式放置以便不干扰机器人的运动。对于这些事实，该架构已被推广并以某种方式简化，得到一个抽象的实验目的。

19.3 通用系统构架

在下面章节，一个基于特定目标的通用系统构架将被介绍。为展示智能家居环境进化的可能性，某些存在的限制有时候会被忽略。在图 19-1 中，描绘的是一种通用系统构架的整体结构，尤其针对的是老年人和残障人士的需要。在我们的例子中，中心利益点是应该转变成一个智能家居的公寓。因此，这种公寓配备了许多技术系统。这些系统需要互相连接尽可能以清楚的方式支持用户。无线网络技术应该

用来连接各种著名的概念，例如安全或者家庭娱乐系统，以及所谓的智能对象。这些智能对象是普遍对象，带有小型集成微处理器，通过无线网络进行通信。它们特定配备有各种传感器来观察环境，也可以把它们当做执行器。智能对象通常在连接普适计算时被提及，以描绘无处不在的信息和移植计算机技术进入日常生活为特征。但是直接和遥远的环境，包括邻居、医生或医院、家庭，现有的网络和因特网技术、公共电话网络和家庭自动化系统，也必须纳入一个智能家居环境理念。

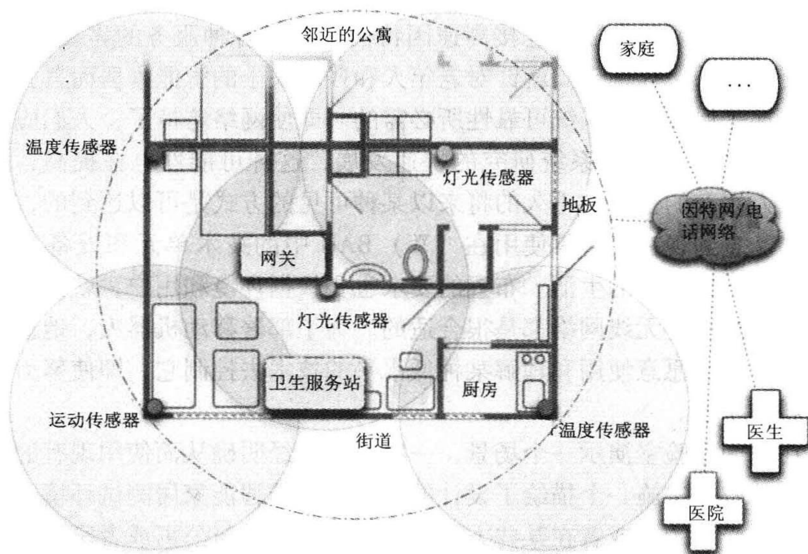


图 19-1　纵观一个普通的系统架构

静态无线多跳网络有时候使用智能对象例如 RFID，可以以一种自组织和自学习的方式控制各种系统。在经济方面，常见的进程，有关本地系统、空调、照明和加热，使他们可以以更智能的方式一起工作。一种著名的技术是使用传感器和执行器设立反馈循环。因此，传感器与执行器可提供以预先确定的规则为基础的功能以及从一个学习的过程演变的行为。因此环境以一种自组织的方式来控制。举几个例子，加热以及照明和空气调节可以被经济地控制。此外，冰箱可以保存许多内容来自动整理直接从超市购买的耗尽食物，然后再传递到公寓。使用 RFID 技术的智能地毯是安全系统的一部分。不过，他们可能是监视系统的一部分，以确保病人还活着，也捕捉运动和其他行为的模式。

除了正常或植入的传感器或者传感器膏药，例如在德国弗劳恩霍夫可靠性和微集成[17]的研究所研制的“智能服装”，即可穿戴电脑和其他智能对象，可以用来留意一起生活的病人。如果可能的话，邻居也可能参与即时的帮助。在不寻常的生命体征情况下，应急系统可以直接通知邻居，使用具有蓝牙功能没有额外费用的设

备。一个可靠的应急系统也必须在紧急的情况下通知医生和医院。为了这些目的，现有的网络，如因特网或电话网，也可使用。该系统可以依靠一些第三方机制。传呼机或者移动设备可能会被称为公共电话网络或直接连接的紧急通知系统，可通过互联网使用，这取决于现有的紧急通知系统提供的特定界面。

随着现在公寓连接到了外部世界，频谱的可能性是急剧扩大。公寓的中心卫生保健站可以收集并提供信息给用户、医疗助理甚至家庭。通信系统也可以从标准的电话到语音/视频改进。因此，医生和病人可以在视频会议中互相讨论，而不用在现实中见面。

显而易见，各种混杂的网络必须被合并。因此，网络间需要一种或者多种网关。传感器网络可以使用蓝牙技术进行通信，同时其余的系统使用使能的 WLAN。由于可以提供多个网关，第一个网关可以连接蓝牙传感器到 WLAN 内部，同时第二个控制进来和出去的电话以及互联网服务。移动机器人对于智能家居环境是另一个有用的延伸，因为它们能够提供不同的服务。上述在传感器网络和无线局域网之间网关的作用例如可以通过一个移动机器人覆盖。一方面，它们可以用来支持和维护网络，增强网络的可靠性；另一方面，观察环境和居民，使用 RFID 技术定位经常丢失对象（FLO）。

无源 RFID 标签的使用消除了电池的局限，而它是传感器网络的一个最具有挑战性的问题。不可能找到一个对象尽管电池是放电的，不管对象是否配备了定位系统。无源 RFID 标签不依靠外部能量供应，它们通常都很小，几乎每个对象都可以配备这些标签。此外，移动机器人的导航援助和分布式协调可以通过植入环境底层的可写 RFID 标签来提供。例如，制造商福维克正在生产集成 RFID 标签的地毯，可以被看做是建立基于超分布式 RFID 标签基础设施的智能环境的基础[10,31]。

19.4 实施

由于测试环境建立在了我们实验室，该构架的具体映射必须在可能的边界之内。在我们实验室，BT 节点传感器节点和称作 Robertino 的移动机器人是可用的。随着有效地实施和来自前期工作的技术基础，考虑到下面的一个章节中指出的局限性，这种构架被映射到了一种实验室装置，将在图 19-2 中描述，它展现了研究环境。可以大致分为以下子网：蓝牙、屋檐局域网、射频识别、因特网（TCP/IP）和电话网。

关于构架，提及传感器网络来进行部署的 BT 节点传感器节点互联和散布网络中的传感器数据值，这通过蓝牙网络来实施。私人无线局域网已经使用无线接入点在实验室建立起来。因此，从一个工作地点或者网络内的其他系统连接到移动机器人是可能的。接入点连接到互联网和大学网络。网关已经建立了专门的服务器，它被 ROSES 方案用于许多目的。它通过大学网络和 Web 服务提供互联网连接。RFID

系统不是一个常识性的网络，但是它提供从 RFID 标签到 RFID 读卡器传输数据的可能性。

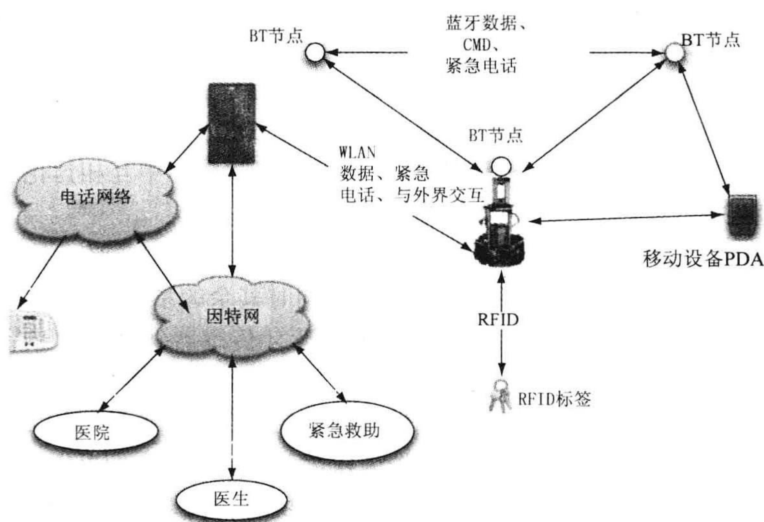


图 19-2 实验装置架构

19.4.1 无线传感器网络

这种无线传感器网络通过一个通用的蓝牙网络以 BT 节点互相连接为基础。关于蓝牙通信层的最上面，一种基于特定按需距离矢量（AODV）的简单多跳协议已经由 BT 核心 API 实施。在多跳协议内可以发射信息给传感器节点，即使它们不是直接互相连接的。那个事实分别简化了覆盖整个实验室或者带有设计的传感器网络的智能家居居住空间的问题。这意味着传感器节点可以放在任何有最大需求的地方。因此，已经解决了有关普通蓝牙通信的有线解决方案和限制性问题。除此之外，使用蓝牙建立传感器网络可以集成使能的蓝牙移动设备。

由于 BT 节点应该作为传感器和执行器节点来操作，一个传感器可以连接到 BT 节点提供的 I/O 线上。存在两种可能的方法来连接传感器或者传感器板到 I/O 线，使用像图 19-3b 中显示的 usbprog rev2 板的连接器或者像图 19-3a 和图 19-3c 中描述的使用两个其他的在 BT 节点上的连接器进行直接连接。传感器可以直接插入到连接器引脚中。这里使用的是型号为 TSL252R 的光传感器。第一种方法中，传感器已经安装在了一个简单的传感器板上来开发一种恰当可靠但简单的解决方法。在图 19-3d 中显示的是电路框图。

第二种方法中，推荐的来自 BT 节点开发网站和相关软件的 BT sense v1.1a 传感器板已经被使用（见图 19-3）。BT sense 板是一种简单的传感器板，被 J2 连接器连接到 BT 节点。它设计于普适计算研究所 ETH Zurich，用于支持学生的教育。板

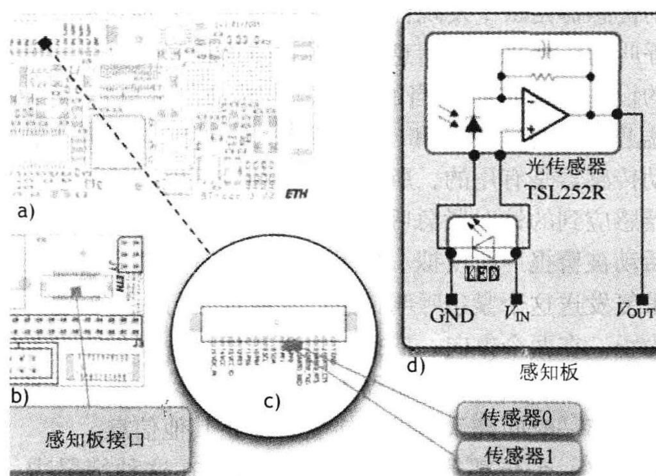


图 19-3　BT 节点插头连接—感知板电路

的大小（2×4cm）允许固定到 BT 节点的一边[9]。

它提供了连接的可能性：

1）TC74 温度传感器（数字，I^2C）；

2）TSL250R/TSL251R/TSL252R 光传感器（模拟）；

3）AMN1 无源红外线运动传感器（数字，逻辑电平）；

4）7BB－12－9 报警器；

5）可选择的 I^2C 数字传感器；

6）一个可选择的外部模拟传感器。

就智能家居传感器网络而言，设置在多跳路由协议顶层的简单协议已经被开发出来。它包含三种类型的信息，根据需要划分成子类。因此如果需要，不同的信息类型可能通过不同的协议控制。主要的信息类型和他们的细分显示在图19-4 中。

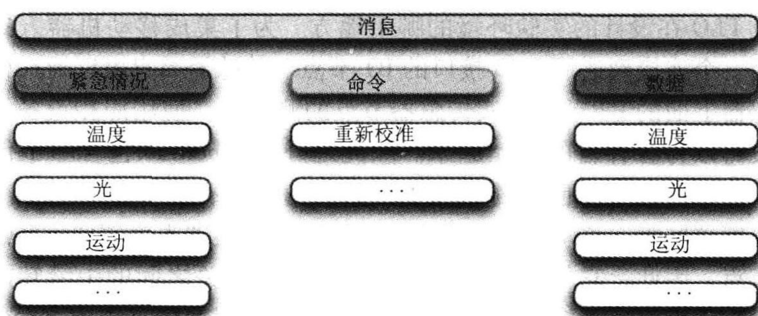


图 19-4　智能家居传感器协议—消息类型

如果覆盖环境的传感器网络中发现任何罕见的参数，发现这种异常现象的BT节点发起一个紧急呼叫的广播来散布信息。如果它使用相同的协议，那么任何集成到蓝牙传感器网络的设备都能够以恰当的方式产生响应。紧急呼叫的信息分为“光传感器信息”、“温度传感器信息”和“运动传感器信息”，并可选择扩展。这里只有光、温度和运动传感器是有用的，所以特定的细分可以满足需求。

由光传感器感应到的针对紧急呼叫的恰当响应是打开/关闭灯光。关于实验装置，执行器的活动被警报声所模拟。如果发起人没有在某个时间限制内收到确认的消息，它将会重复发送这个紧急呼叫信息。完成这个行为后，它持续到执行器确认收到紧急呼叫为止。在两个方向，广播信息都被使用，因为来自其中一个节点的确认信息应该阻止其他节点或者移动机器人的回应。所以，如果有紧急呼叫，最迅速的解决方法是通过第一时间确认请求来接受和阻止其他信息。

命令呼叫来开启一个传感器节点的远程进程。在这种情况中，就如上面提到的，校准的命令已经被执行，通过闪烁LED灯模拟一个确认的进程。命令呼叫也可以被接收传感器节点确认。这种情况，因为没有涉及其他设备，所以信息被传送到起始点。

数据呼叫是被断断续续发送来维护环境状况的消息。收集的数据可以在智能家居内部和外部分析或者仅仅显示给用户。名字“call”被选择来保持命名的一致性。事实上，这是一个简单的未确认的信息。

19.4.2 移动机器人

我们正在使用移动机器人平台Robertino，我们已经为它开发了许多硬件扩展（包括下面描述的RFID读卡器板）。基础软件的功能由新发明的构架Robrain提供[16]。我们在许多学生计划中开展Robrain。这也包括许多插件程序，例如RFID连接器“rfidReader”，定位和映射单元“PathFinder”和摄影机延伸“VDCUnit”。智能家居插件程序混合提供的解决方法来执行一个行为插件程序，它定义了智能家居场景内的移动机器人行为。Robertino必须要集成到传感器网络，能够通过传递它们来定位FLO在设计的实验环境的哪个地方。为了集成移动机器人到传感器网络，会使用一个连接到通用USB接口的BT节点。传感器节点可以通过使用机器人和相应传感器上面的BT节点证实传感器节点的测量值。使用照相设备，图片和小的视频片段可以为各种目的存储起来。除此之外，Robertino包含一个传感器网络基站和描绘外围世界与传感器网络之间的接口。

关于传感器网络，移动机器人有两个作用。首先机器人为邻居传感器节点建立数据汇聚节点。因此，它们作为传感器节点和中心网关之间的中继转播设备。此外，移动机器人应该负责维护和管理有关验证和替代以及传感器节点和数据存储校准。因此，如果机器人发现许多未确认的紧急呼叫，它将检查传感器节点。如果传感器节点显示有怀疑的行为，移动机器人通过开始一个命令呼叫指示BT节点重新

校准。在这两种情况下，Robertino 需要拍张当前形式的图片，并通过蓝牙短信或者使用互联网（没有蓝牙功能的设备可用）通知下一级实例用户。

机器人应该根据预定的路径观察实验室。为了管理传感器网络和观察环境，环境范围内机器人的导航必须在环境中制定出一定的坐标定位。因此，使用了路径搜索插件。关于插件的能力，它是预先确定要找到一个从给定出发点到某个目标的最佳路径。如果机器人基于上面提到的原因决定接触传感器节点，该插件找到的是环境指定地图的最佳路径。

由于没有进一步的定位方法被部署，安装的 BT 节点必须标明在环境地图上。因此，许多已安装的 BT 节点设置用地图上的一个区域来分配传感器节点的地址。此外，移动机器人现在可以使用路径搜索插件所提供的动态模式更新给定的地图。穿越了一个预定义轨道上的机器人能够使用安装在 Robertino 上的 RFID 读卡器系统定位 FLO。考虑到使用在地图搜索插件的地图坐标，如果 FLO 被发现了，该对象的 ID 将被分配到当前地点。

19.4.3　射频识别

在讨论 RFID 模块实施之前，我们介绍一些与我们的场景相关的 RFID 基础。如图 19-5 所示，RFID 系统主要包括三部分：RFID 标签、读卡器和信息服务主机。读卡器使用天线发射信号，也为无源 RFID 标签提供必要的能量辨认传播和应答疑问。这个信号由读卡器产生响应申请。在我们的场景中，我们定期地调查环境，当 RFID 标签在传输范围之内时，它把 EPC 传输给读卡器。读卡器获得的数据然后发送给中心计算机。

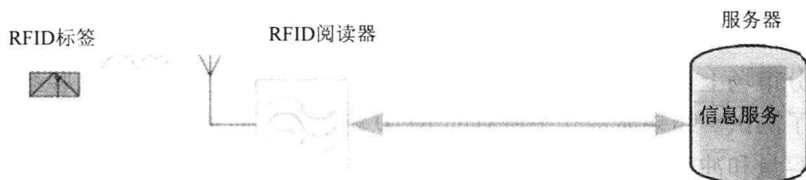


图 19-5　RFID 传输系统

市场上存在几种 RFID 标签：无源、有源和半有源。无源标签没有内部能源，而是使用天线上的低功率信号，它来自读卡器去提供能量和发送应答。无源标签可以在几厘米到几米的范围与读卡器通信。另一方面，有源 RFID 标签有电池（预期电池寿命达到了 10 年），使它们有一个更大的传送范围，甚至更大的存储能力。但是它们更大且更贵。半有源（或者有时候叫半无源）标签使用内部电池为电路供电，但是它们依靠读卡器进行信号传输。

为了测试的目的，我们把型号为 TLB-12-AA 的 RFID 读卡器（图 19-6 右图描

述）连接 Robertino 机器人较低的一边。为了选择合适的 RFID 读卡器，几个问题应该考虑。针对我们的场景，我们决定使用无源 RFID 标签。因此，我们处理的范围大约是 10cm。因为我们将要把读卡器附在一个机器人上，所以它应该有一个低功耗的设备。此外，大小应该是紧密填满的，价格应该是合理的。在图 19-6 左图中，展示了改装的 Robertino 以及在我们智能家居中使用的典型 RFID 标签。

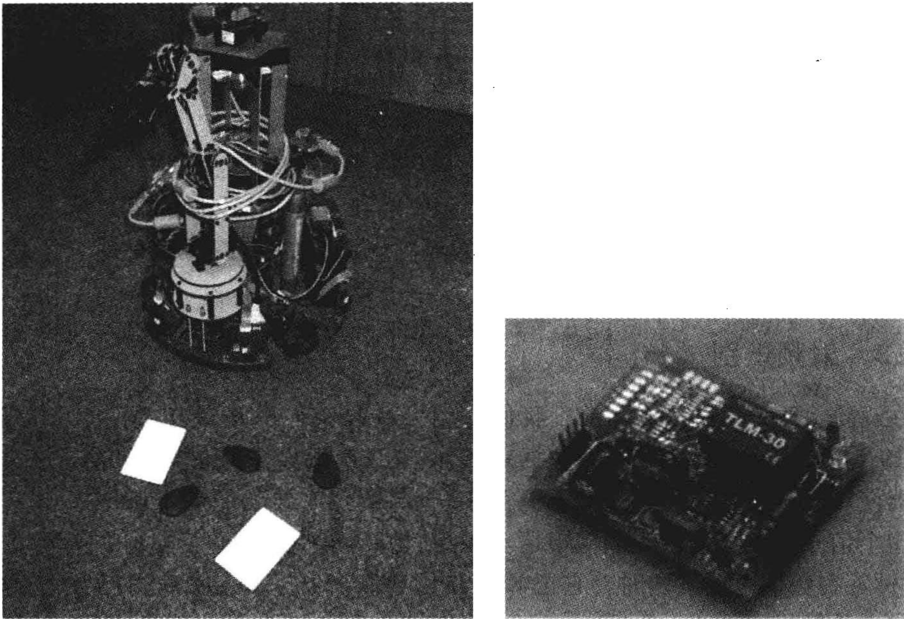


图 19-6 装有 RFID 阅读器（钳爪下）的 Robertino 机器人和一些典型的被动 RFID 标签（左）以及被选择的 RFID 阅读器模块 TLB—12—AA（右）

大部分市场上典型的 RFID 标签列于表 19-1 中。我们买了几个这些系统并在实验室进行了测试。结果显示几乎所有的 RFID 标签都可以使用不影响应用。主要的区别在于容量和冲突检测和安全测试的有效性。

表 19-1 在大市场中可用的典型的被动 RFID 标签的比较

收发器	频率	ID	EEPROM	防碰撞	加密
Unique	135kHz	40bit	N/A	N/A	N/A
Hitag - 1	135kHz	32bit	2048bit	是	是
Hitag - S	135kHz	32bit	32/256/2048bit	是	是
Hitag - 2	135kHz	32bit	256bit	是	是
Mifair	13.56MHz	32bit	1/4KB	是	是
EPC Gen - 2	860 ~ 960MHz	64 ~ 96bit	64 ~ 96bit	是	是
μ - Chip	2.45GHz	128bit	N/A	N/A	N/A

RFID 读卡器由我们的机器人控制系统 Robrain 叫做“rfidReader”的特殊插件控制。读卡器的输出存储在像 XML 的编码文件夹中，一个如图 19-7 所示。它包含读卡器单一呼叫样本输出。第一行的 OK 表示有序的（方法）呼叫已被成功执行。第二行的 RETCODE 列出了错误代码。被发现的 RFID 标签是 UNIQUE 型号，它包含 1234567890 的 EPC。

<RETSATUS> OK</RETSTATUS>
<RETCODE>200</RETCODE>
<NAME>get Unique Nonblocking</NAME>
<INFO>UNIQUE</INFO>
<OUTPUT>1234567890</OUTPUT>

图 19-7 像 XML 编码的 REID 阅读器输出

19.4.4 网关/手机

网关是家庭网络和外部世界的桥梁。Apache 网络服务器使传感器网络收集和 Robertino 储存的数据在智能家居外部可用。紧随着构架的设计，语音/视频呼叫的方法被实现了。基于许多公共的和商用的方法已经进化的事实，没有必要再测试这些系统。因此，这个系统不会在我们的实验装置中实施。提到的方法即使是改进过的可以很好地满足要求，尤其在可靠性方面的改进是理想的。

为了显示移动机器人和移动设备之间通信的可能性，Robertino 应该能够发射 SMS 信息给手机。我们的方法中，信息通过使用 BT 节点上使用的蓝牙技术进行传输。因此，不需要更多的硬件，此外成本也不会提高。

19.5 实例

这个章节中描述的实例用来测试和评估介绍过的构架。建立实例场景来显示特定解决方法的相互作用。图 19-8 展示的图片是实验的装置。首先，一个区域已经被限定来描绘一个智能家居环境。在我们的实例中，它包含一个小型的房间，包括一个生活区和一个厨房。区域的周围是用木质平板框架起来的。一个传感器网络连接三个 BT 节点，每个节点配备 BT sense v1.1a 传感器板和相应的传感器。每个节点配置使用一个传感器。因此，一个光、温度和运动传感器节点是有效的。每个传感器节点每隔 30s 即时的发送测量的传感器数据。时间间隔已经通过估计来限制演示大约 5 ~ 10min。

如果传感器的值超出了范围，节点就会产生紧急呼叫。对这些值进行预定义以满足实验室的需要。如果光照的值下降 $300\times10^{-6}\mathrm{W/cm^2}$ 或超过 $800\times10^{-6}\mathrm{W/cm^2}$，或者温度下降到 15℃或者超过 35℃，或者运动传感器在最后时刻没有监测到任何运动，那么这个值超出了范围。第四个节点被配置为一个简单的执行器，来演示一个反馈回路。在我们的场景中，如果光照的值超出了范围，BT 节点起到一个开关的作用。这通过使用 BT sense 传感器板上的警报器发出连续的嘟嘟声进行模拟。

运动传感器用来显示移动机器人的行为。如果传感器发送五个以上未答复的紧

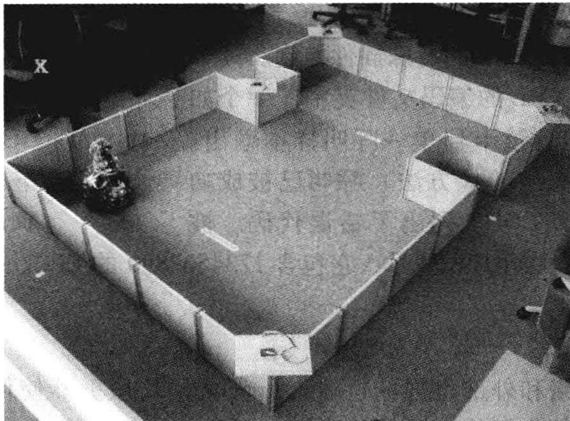


图 19-8 实验装置

急呼叫，机器人就会停止观察生活区，发送一个答复（确定呼叫），直接移动到节点。它为实际环境进行拍照并且通过蓝牙发送 SMS 信息给手机。然后 Robertino 返回到之前停止的地点并且继续观察生活区。观察路径、机器人期望的移动和 BT 节点的位置如图 19-9 所示。在 80 端口上提供 Web 服务的主机生效，执行一个 CGI 脚本。主机连接到实验室特有的 WLAN 以及 Robertino。

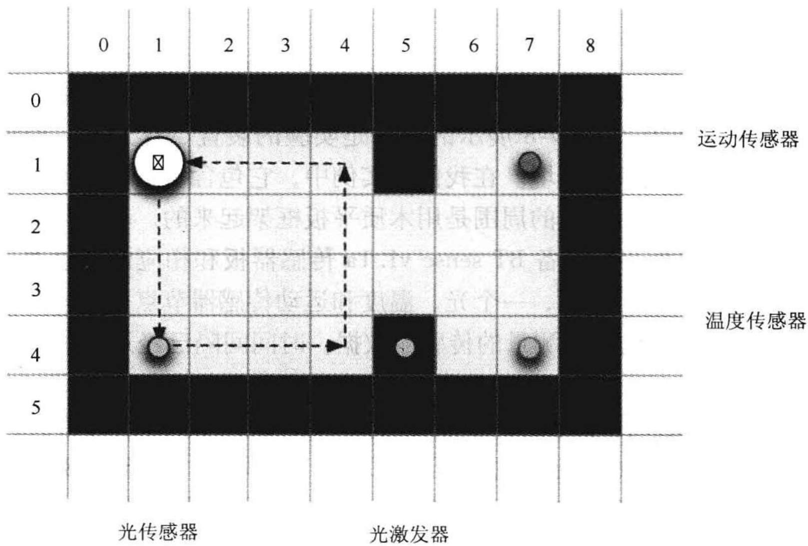


图 19-9 网页快照

一个动态生成的 Web 页面提供了所有收集的数据和相应的情节，如图 19-10 所示。Web 页面包含显示于左手边的状态信息。此外，从移动机器人上下载的最

新数据项目被列了出来。事实上，这个清单描绘了对传感器网络成功的测量。传感器数据在 Web 页面的底端更加详细的展示。这里所有的温度、光和运动监测值都列了出来。为了进一步给用户提供数据信息，细节可以通过所有这些测量值创建。

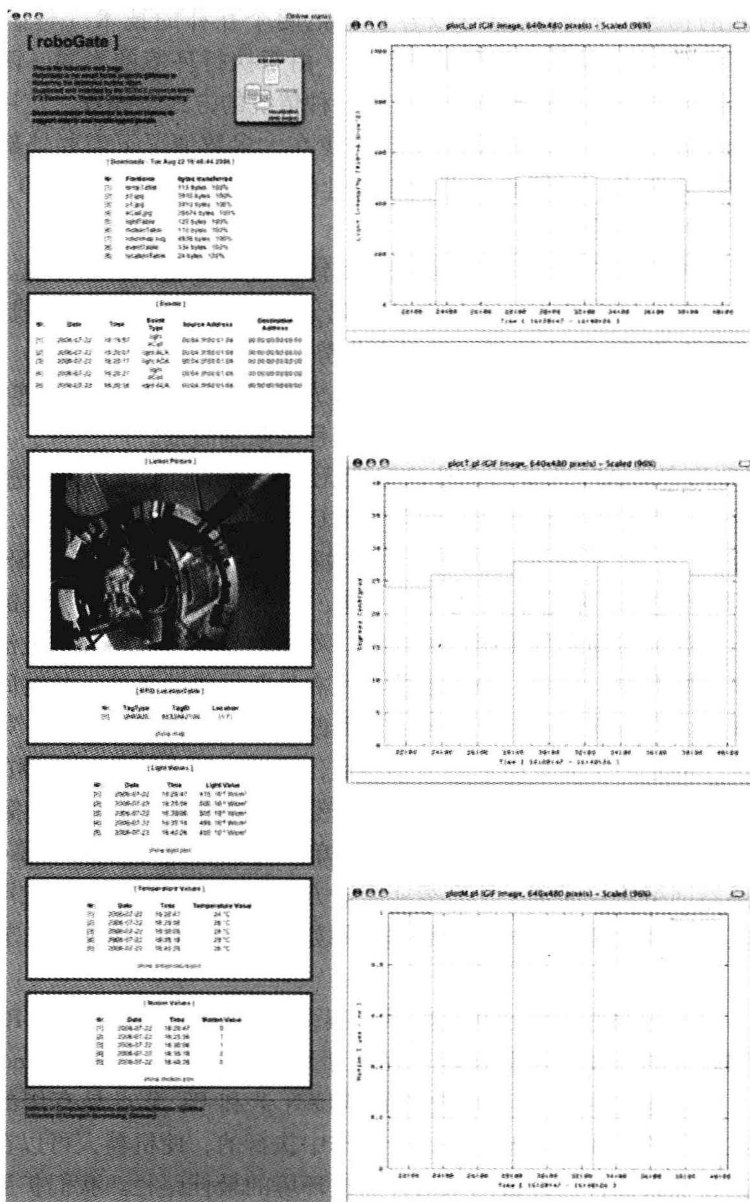


图 19-10　实验装置—图解

此外，从RFID标签收集的信息包含在Web页面中。中间描绘的是RFID位置表。可以看出，单个RFID对象最近已经被定位并且位置可以描绘到图上。

19.6 实施体验

使用成熟的构架，我们期望可以容易地把两个互补的技术（传感器网络和RFID标签）结合起来。从软件工程的观点看，由于执行环境的不同，两个系统需要一个完全不同的方法。传感器节点通常以一种传统的方式来编程，应用软件的开发、安装以及在节点上运行。相反，RFID至少在大多数情况下需要外部决策和提供唯一有限的计算能力，尽管它们有精心设计的子存储系统。与RFID系统相比，传感器网络传送的距离更远。

为了开发一种集成的方法，两个系统的具体能力需要被识别和开发。而且，在这种环境中运行的应用需要为各部分清楚地划分模块。如果这个进程成功地实现，集成传感器网络和RFID的方法提供了很多优势。

在我们的实例中，我们用RFID技术把传感器网络和移动机器人连接了起来。在下面，我们总结主要的问题，它们是实施进程的限制因素。

BT节点和相应的API非常适合快速设置无线传感器网络和展现传感器网络应用的原型。这些节点配备有射频接口和各种I/O线。考虑到它们的大小，节点提供了惊人的计算能力。虽然很容易为BT节点编程，但是随着软件在ANSIC中被执行，API被不断改进来扩展它们的能力。此外，一个非常好的不断扩展的教程是有用的，它展示了如何实施和使用提供的功能。不幸的是，众所周知的电池和内存限制在某些方面仍然存在。所以，节点不能在没有外部供能的条件下长时间运行。有关我们的实施，内存方面没有什么困难制约着改进。关于传感器硬件，第一种方法是我们自己设计的是一个非常简单的传感器板，而第二种方法使用推荐的BT sense v1.1传感器板。两种方法都可以很好地解决感知环境的任务。然而，第二种方法更具有兼容性。BT nut API提供的方法解决了单一传感器在使用BT sense传感器板的问题，传感器网络协议涵盖了所有智能家居方案需要的必要信息。它可以针对紧急情况产生反应以及收集普通的环境数据，例如通过开发执行某些BT节点的命令的可能性。

Robertino已经被很好地设计并且提供了很多的可能性应用。随着ROSES计划内新的硬件和软件的开发，这些技能正在不断地增加。不难把Robertino集成到使用的网络中，因为连接到USB串行设备的WLAN卡和BT节点是有用的。这个插件，不管是在之前的工作开发的还是这个工作中执行的，使机器人可以在推荐的方式下工作。路径探索插件已经被用来估计通过环境的最佳路径。我们扩展了这种插件来访问机器人的当前位置。RFID读卡器和照相机解决的任务没有任何重大问题。然而，只有类似于卡的RFID标签可以使用，因为读卡器的读取范围被限制在大约

10cm，并且机器人必须能够顺利通过标签。

关于 SANET 和 RFID 技术的结合，必须要说的是许多小问题，诸如典型的 RFID 读卡器读取距离近（至少如果无源 RFID 标签被连接（考虑）），需要对移动机器人进行非常精确地移动控制。诸如读卡器到机器人恰当的物理连接和由于机械方法的其他问题的非方法问题依然主导着智能家居演示装置的发展和部署进程。如果商业问题（可能会更加昂贵）可以被定义并使用在智能家居应用不断增长的市场中，很显然这将更加容易处理。

19.7 结论

随着对普适计算方面兴趣的不断增加，需要（提供了）各种各样新的、小的并且易于使用的技术。使用于控制单元或者移动机器人的智能对象、传感器节点、执行器和其他小而对我们生活有很大帮助技术已经变得越来越有用越来越实惠。考虑到快速增长的研究和发展进程，很快这些对象将变得像现在的手机和 PDA 一样普遍。关于医疗健康，让病人居住在自己熟悉的环境接收观察和建议，而不是在医院里待数月甚至几年将会变得可能。医疗和老年人护理在紧急情况下将变得更加容易和安全，因为及时的帮助是很有用的。考虑到所有现实，传感器/执行器网络会很适合在智能家居环境中帮助老人和残障人士。

实验性智能家居装置描绘了研究智能家居环境方面一种很好的方法，可以说许多事情已经被抽象到某种程度并且许多问题依然需要解决。关于与智能家居场景相连接的成熟系统的可靠性，尽管一些部件不仅自身可以很好地工作而且也能与其他系统很好的结合，但已经注意到某些不足。另一个关于系统可靠性的问题关系到无连接的蓝牙多跳网络，因为它没有提供任何数据传输可靠传输层机制。

面临的挑战是设计一个智能家居环境尤其是帮助老年人和残障人士。在任何情况下，智能家居技术融合传感器和执行器异构网络、自动化系统、移动机器人和互联网能够补偿某种限制，并可能帮助用户活得更长，在生活的许多方面更加独立。

参考文献

1. Center for Future Health—Smart Medical Home Research Laboratory, 2001–2005. Online; http://www.futurehealth.rochester.edu/smart_home/.
2. INGA e.V.—Innovationsnetzwerk Gebäudeautomation, 2003. [Online; http://www.inga.de/].
3. T-Com Haus—Intelligenter Wohnkomfort, 2006. Online; http://www.t-com-haus.de/.
4. I. F. Akyildiz and I. H. Kasimoglu. Wireless sensor and actor networks: Research challenges. *Elsevier Ad Hoc Network Journal*, 2:351–367, October 2004.
5. A. Awad, T. Frunzke, and F. Dressler. Adaptive distance estimation and localiza-

tion in WSN using RSSI measures. In *10th EUROMICRO Conference on Digital System Design—Architectures, Methods and Tools (DSD 2007)*, pp. 471–478, Lübeck, Germany, August 2007, IEEE.

6. M. A. Batalin and G. S. Sukhatme. Coverage, exploration and deployment by a mobile robot and communication network. In *International Workshop on Information Processing in Sensor Networks*, pp. 376–391, Palo Alto, April 2003.
7. Bauland GmbH. Vision Wohnen, 2005. Online; http://www.visionwohnen.de/.
8. O. Beisheim. Futurelife—Smart home Beisheim-Gruppe metro. Online; www.futurelife.ch.
9. J. Beutel, M. Dyer, O. Kasten, M. Ringwald, and K. Römer. BTsense V1.1a, 2006. Online; http://www.btnode.ethz.ch/.
10. J. Bohn and F. Mattern. Super-distributed RFID tag infrastructures. In *2nd European Symposium on Ambient Intelligence (EUSAI 2004)*, pp. 1–12, Eindhoven, the Netherlands, 2004. Springer.
11. Continental Automated Buildings Association. OnStar, 2002. Online; http://www.internethomealliance.com/pilots_projects/family/onstar_at_home/.
12. D. Culler, D. Estrin, and M. B. Srivastava. Overview of sensor networks. *Computer*, 37(8):41–49, August 2004.
13. S. Dengler, A. Awad, and F. Dressler. Sensor/actuator networks in smart homes for supporting elderly and handicapped people. In *21st IEEE International Conference on Advanced Information Networking and Applications (AINA-07): First International Workshop on Smart Homes for Tele-Health (SmarTel'07)*, Vol. II, pp. 863–868, Niagara Falls, Canada, May 2007, IEEE.
14. M. K. Divyan and K. Kwangjo. Security for RFID-based applications in smart home environment. In *2007 Symposium on Cryptography and Information Security (SCIS 2007)*, Sasebo, Japan, January 2007.
15. F. Dressler, *Self-Organization in Sensor and Actor Networks*, John Wiley & Sons, New York, December 2007.
16. F. Dressler and M. Ipek. An extensible system architecture for cooperative mobile robots. Technical Report 06/06, University of Erlangen, Department of Computer Science 7, December 2006.
17. R. Dünkler, Body Area Network—Aufbau- und Verbindungstechnik (IZM), 2002. Online; accessed September 5, 2006.
18. H. Elzabadani, A. Helal, B. Abdulrazak, and E. Jansen. Self-sensing spaces: Smart plugs for smart environments. In *From Smart Homes to Smart Care: 3rd International Conference on Smart Homes and Health Telematics (ICOST 2005)*. IOS Press, Sherbrooke, Québc, Canada, 2005.
19. W. Friesdorf. SENTHA—Seniorengerechte Technik im Alltag/Technik im Haushalt zur Unterstützung der selbständigen Lebensführung älterer Menschen, September 1997. Online; http://www.sentha.tu-berlin.de/.
20. B. P. Gerkey, R. T. Vaughan, and A. Howard. The player/stage project: Tools for multi-robot and distributed sensor networks. In *International Conference on Advanced Robotics*, pp. 317–323, Coimbra, Portugal, July 2003.
21. S. Glubrecht, K. Greiner, D. Wichmann, and E. Steidl. Haus der Gegenwart-Wir wollen wissen, wie wir wohnen, 2006. Online; http://www.haus-der-gegenwart.de.

22. G. Goldacker. LIVEfutura. Online; http://www.livefutura.de/.
23. S. Helal, W. Mann, H. El-Zabadani, J. King, Y. Kaddoura, and E. Jansen. The gator tech smart house: A programmable pervasive space. *Computer*, 38(3):50–60, March 2005.
24. N. B. Priyantha, H. Balakrishnan, E. D. Demaine, and S. Teller. Mobile-assisted localization in wireless sensor networks. In *24th IEEE Conference on Computer Communications (IEEE INFOCOM 2005)*, Miami, FL, March 2005.
25. H. Ruser. Technische Universität der Bundeswehr München—Smart HOME, 2002. Online; http://smarthome.et.unibw-muenchen.de/de/.
26. K. Sakamura. TRON—Intelligent House, 1998. Online; http://tronweb.super-nova.co.jp/tronintlhouse.html.
27. K. Scherer. inHaus Duisburg—Innovationszentrum—Intelligente Raum- und Gebäudesysteme—Phase1, April 1995. Online; http://www.inhaus-duisburg.de/.
28. E. M. Tapia, S. Intille, and K. Larson. Activity recognition in the home using simple and ubiquitous sensors. In *Pervasive Computing*. Springer, Berlin, Heidelberg, 2004.
29. W. Trumler, F. Bagci, J. Petzold, and T. Ungerer. Smart doorplate. *Personal and Ubiquitous Computing*, 7(3):221–226, 2003.
30. VeriSign. The EPCglobal network: Enhancing the supply chain, 2005. White Paper.
31. Vorwerk. Vorwerk is presenting the first carpet containing integrated RFID technology, 2005. Online; http://www.vorwerk-teppich.de/sc/vorwerk/rfid_en.html.
32. Wikipedia. Intelligentes Haus—Wikipedia, The Free Encyclopedia, 2006. Online; http://de.wikipedia.org/w/index.php?title=Intelligentes_Haus&oldid=1917438.

第20章　应用于卫生保健系统的RFID与无线传感器网络的集成

射频识别（RFID）和传感器网络这两种无线技术很快将会成为日常生活必不可少的一部分。它们在普适计算方面有许多潜在的应用。一方面，RFID在工业技术和应用方面有了很快的发展；另一方面，传感器网络技术在其研究领域中已经引起了巨大的关注。这一章节试着架起工业与学术界焦点的桥梁。我们呈现了RFID和传感器网络的研究在卫生保健方面的应用。随着老龄化人口的增长，这是一个极其重要的领域，将变成一个普遍的、全世界的现象以及问题。

20.1　概述

RFID和传感器网络这两个无线技术未来具有无限的潜力。随着工业见证了RFID技术的发展和应用的快速增长，以及网络研究团体在传感器网络方面投入了巨大的努力，这两个团体将会从互相学习中获利。

无线传感器网络最近对科学家和研究人员、军队和政府官员以及商业团体提出了许多令人兴奋地挑战。这很大程度上是由于其提供的无限的潜在应用。一些目前著名的应用包括由很小的无线微电动机械传感器组成的“智能节点”网络，能够追踪和报道从军队操作方面敌人运动到房间里寒冷的地方几乎所有的事情。除此之外，先进的无线传感器也能够发现制造方面的缺陷，追踪肥料从农田流失到湖里或者监测医院里病人的移动情况。

尽管传感器网络能够感知许多不同类型的对象，然而它也有自己的局限。它通常感知对象的物理、化学或者生物特性（例如温度、光照、声音或者移动）。没有这些可感知特性，对象将很难被感知。通过把RFID作为一个特殊的传感器应用于传感器网络，任何带有RFID标签的对象都可以很容易地被“感知”（即识别），因此克服了传感器网络的一个重要缺点。这种结合被认为是在RFID和传感器网络方面下一个提高——不仅在技术方面，也在于打开了通向其他过去从来没有存在而又有应用的可能性的大门。

许多最近的研究表明人口老龄化是一个全世界的现象。在美国65岁以上的人口将会从1975年的1060万增长到2025年的1820万，增长了72%，而总体人口的增长只有大约60%。老龄化趋势不仅制约着美国，事实上也是个全球性的现象。从1990年的3570万到2025年的7610万，世界上大于65岁的人口将增长两倍多。长寿已经产生了年龄相关的贵重代价，例如残疾、疾病和卫生保健花费。

为追求集成RFID到传感器网络技术的巨大潜力，同时由于老龄化人口卫生保健的需要，这一章节我们的研究目标是卫生保健领域。注意我们在卫生保健方面的研究建议和方案将利用这两种技术。因为焦点是两种技术的结合，这一章节中我们不包含那些只使用一种单一技术（传感器网络或者RFID）的研究。

这一章组织如下：20.2节描述三种目前在医院中RFID和传感器网络使用的建议，20.3节描述三种医院外卫生保健方面的应用，20.4节提出和说明一种集成RFID和传感器网络的卫生保健平台。最后20.5节对这一章进行总结。

20.2 智能医院使用RFID和传感器网络的调查建议

这一节中，我们首先描述一种应用，分析和管理医院人员的供应和需求[7]。接着的是保持和追踪非常敏感的医疗供应[7]。最后，给出了一种关于建立普遍感知医院的建议。

20.2.1 医院人员流动供应和需求管理分析

利用RFID独特的特性能够精确地识别一个人，Xiao等人研究了关于医院人员的供应和需求管理，诸如医生、护士和病人[20]。这些人身上都贴着RFID标签，以便他们的流动能清楚地被识别，并且供应和需求的瓶颈可以被调查和提高。

上面的应用中，RFID标签作为塑料带被戴在一个人的腕关节上。与每个病人唯一的ID相联系的是病人的名字、出生日期、性别以及详细的医疗记录次序信息等，都存储在一个远处的数据库里（DB）。对于医生和护士，RFID标签可以被嵌入到他们的访问ID，这经常使用于进入房间的各个层次。

这些唯一的ID会被RFID读卡器监测到，读卡器可以固定在每个房间里或者随无线连接的片式个人电脑（PC）移动。这些RFID读卡器通过无线网络连接到DB服务器，这里存储着这些特有的ID和他们相联系的信息。

医生和护士的活动可以在重要的地方进行分析，那里资源相当稀缺和重要，例如急诊室。这可以通过记录他们的流动来实现，包括每个人在某个病人身上或者等待某种供应花费的时间。因此，他们的活动可以被仔细地分析，并且可以识别出瓶颈。这将会提高人员流动和资源管理的有效性。

另一方面，也可以分析病人访问和服务。这可以通过记录病人流动实现，例如病人花费在某个治疗时期的时间，包括挂号、各种等待、测试和实验。这会潜在地提高病人的护理和治疗。

这些应用中，使用的RFID主要是医院里的人可带的RFID标签和相关RFID读卡器。传感器网络可以加入到无线基础设施中来提高连通性和进行数据收集。

20.2.2 追踪重要的和非常敏感的医疗/生活供应

尽管 RFID 可以精确地识别一个人或者一个对象，就如上节的研究中所展现的，但是先进的传感器可以感知很多复杂的对象特性。Kim 等人发明并客观地论证了一种系统，它能够通过监测血袋的温度追踪它们的位置[7]。用于这个方案的特有传感器和 RFID 设备已经被仔细的描述过。

血液是医院里非常宝贵的生命必需品。它对温度非常敏感，并且应该存储在温度稳定的环境中。事实上，它必须保存在 2 ~ 6℃ 的环境中，以便保持它的质量和稳定性。如果把它放在室温里，即使只有 30min，红细胞将会溶血并且血液的寿命会剧烈减短。

医院的血库管理着每个血袋——从血液捐赠者那接收到它的时间开始到把它给血液接受者输血为止。因此，保持和监测血袋的温度和位置是至关重要的，以便于保持它的温度和可用性，迅速正确地给正确的病人输血。使用 RFID 和传感器网络技术，作者开发了一个提高 3T（时间、温度、追踪）的系统来提高对血液样本的温度管理和阻止病人血液不匹配。这种系统在设计的时间间隔追踪血液袋的移动和位置，监测它的温度变化。它也产生医护人员分享的数据报告。

这种系统使用 Crossbow 技术的 MTS420 传感器来记录血液袋、冰箱和血库的温度。它也用来和 MICAz 节点连接来监测相关的因素，诸如湿度和光加速等。ZigBee 或者 IEEE 802.15.4 无线技术用来传输温度数据给汇聚节点 MIB510CA 传感器，它也附在 MICAz 平台节点上。

对于 RFID 温度传感器，也就是 RFID 标签和 RFID 读卡器，使用来自 KSW Co. 的 TempSens。它们是有源 RFID 标签使用 13.56MHz 带宽。各自有一个可以持续超过 16 个月的内置片式电池和 64 个温度测量数量的 SRAM 存储器。这种系统也包含一个位置追踪系统（LTS），它包含主机、容器、站、操作者、控制器和移动控制器。

系统进程开始于血库的冰箱。在这里面，RFID 和传感器收集温度数据（血袋 ID、冰箱温度等）与外面的汇聚节点进行通信。然后将数据发送给笔记本计算机和 RFID 读卡器。这个数据然后被存储到 DB 服务器中，它连接到医院的信息系统，使医护人员可以通过本地无线网或者 Web 服务来获取信息。

血袋从冰箱里拿出来后，LTS 把它传送到最终确定的地方。LTS 确保可以跟踪这些血袋，医护人员对它们的到达保持着警惕。DB 服务器也允许医护人员查询血袋发放的时间和它确切的位置。

这整个系统已经被测试，它的可用性和有效性已经在韩国首尔 Shin - chon Severance 医院被证实。已经证明血袋传输自动智能系统提高了效率和医护人员的努力。除此之外，这个系统提供了一个更加精确分析数据的历史；因此它减少了人们出错、使用劣质血和丢弃血液的几率。

20.2.3 建立一个普适感知医院

基于以上呈现的两种应用，应该清楚RFID和传感器网络能被用来建立一个智能普适医院。Wu等人提议通过促使RFID和传感器网络技术改变来建立一个智能医院[18]。

类似于前面讨论的两个方案，这个方案中，创始人发现RFID和传感器网络提高精确俘获数据的效率。RFID标签可以用作卫生保健专业人员和病人的ID标记，也可以用于各种目的的医院固定资产中。通过RFID标签和传感器网络的帮助，医生和护士在紧急情况下可以迅速地被追踪和取得联系。

除此之外，RFID和传感器网络技术帮助提高医院的药物管理。例如，它们可以用来追踪药物分配和药物用量的有效性。它们在实验室的应用包括病人样本有效性和提高实验结果精确性，追踪血液供应以及血液匹配（如前面的章节清晰的解释）和控制处理废材料的有效性。

这些应用的一个挑战是RFID和传感器节点产生大量实时或者近似实时的数据。这个数据很难以有效地方式被组织、管理和使用。应对这一挑战，创始人建立了一个系统eWellness，组织和便于管理基于事件概念的传感器数据。eWellness包含一个旁边范例，它组织医院RFID传感器数据和一个三层事件代理，并推理绘制应用领域中事件的真实的模型[18]。

旁边设计基于一个事件的概念，在医院应用的背景下，塑造内在的短暂空间RFID和传感器数据的特性。例如，跟踪一个输液泵在没有批准的情况下离开急诊室时一个事件会被触发。

传送事件数据提供了以实时模式处理大量混杂传感器数据和保持医院里不相关分立事件的关系，以及以短暂、空间和因果环境来联系那些事件的能力。

一个追踪输液泵的原型系统已经在当地医院测试过，使用初步数据显示了在管理有效装备方面的提高。与这种原型相联系的一种事件本体论语言EO随后被提了出来[19]。它发展用来帮助建立应用环境以及推理事件。

20.3 医院外卫生保健使用RFID和传感器网络的调查建议

这一节首先解释了一个给病人提供移动遥测服务的建议，目的是提高远距离诊断和减少医疗紧急情况下危急响应[20]。然后讨论了无线健康监测系统Health Tracker2000，最后，描述了针对老年人的家庭医疗帮助系统的原型。

20.3.1 移动遥测服务

除了提高医院里病人护理的效率和管理，RFID和传感器网络技术也为医院外的病人提供特有的优质卫生护理。Xiao等人提出了一种实时病人监测系统，使用

RFID 和智能传感器来收集病人的重要信息，包括心电图（ECG）、脉搏、底部温度、血液含氧量和食道酸性[20]。除此之外，智能位置追踪功能也包含定位病人，尤其在紧急情况下。

这个系统中，微传感器无扩散的连到病人搜集重要信息，包括心电图、脉搏和底部温度。RFID 标签用来识别和定位病人。传感器和 RFID 标签搜集的数据然后分别发送给传感器头和 RFID 读卡器，接着传送给 PC、私人数字助手（PDA）或者甚至手机。医疗人员然后能够监测病人的重要信息以及无论何时何地进行远距离诊断。这种系统因此提高了远程诊断的精确性并减少了医疗紧急事件的危急答复。除此之外，在医院外它能够潜在提供普遍的卫生保健系统。

20.3.2 无线健康监测系统

通过在医疗领域和婴儿潮一代的老龄化的大量研究，Teaw 等预测美国人的平均寿命将会快速增加[16]。高龄人口最迅速的井喷将会在 2011 ~2030 年间发生。在这 19 年期间，老年人将从总人口的 13% 增加到 22%。由于这个现象，相当多的研究团体已经提出为老年人卫生保健使用 RFID 和传感器网络，这节和下一节将会说明这种方案。

Teaw 等人设计一种无线健康监测系统——健康追踪系统 2000[16]。它由无线传感器网络、RFID 和目前的重要信息监测技术组成，后者在追踪用户位置的同时监测重要信息。

这种系统包含携带病人追踪系统的患者，追踪系统把重要信息发送到基站（BS）。监测的重要信息包括心率/脉搏、血压、呼吸频率和体温。BS 搜集各种数据然后把聚集的数据发送到监测站，监测站可能与一个家庭成员或者卫生保健提供者或者在医院内的人在一起。病人追踪设备通过现在的技术可以实现。一个或者多个病人可以从单个 BS 被监测到。这个系统可以在所有类型的家庭或者设备中安装。RF 波发送病人的重要信号和位置信息给他们的亲属或者医护人员，也会在生命危险的情况下发送警报。

为了精确地测得病人重要的信号，创始人仔细报道了他们选择的各种传感器和 RFID 设备。重要的信号包括温度、脉搏、呼吸率和血液含氧量。对于热量传感器，国家半导体选择 LM92 而不是 LM34，因为它有一个最小的供电电压 2.7V 和 0.3℉ 的分辨率（在 LM34 中为 0.5℉）。对于血氧计，Nonin ipod 用作数字血氧计来检查用户的手指。它可以测量血液含氧量以及脉搏。呼吸率传感器被认为很难发现，创始人还没能够找到适合他们的系统的呼吸率传感器。

健康追踪 2000 系统的追踪部件也是用 RFID。类似于上面提到的系统，在每个人身上都穿有个 RFID 标签，并且有一个 RFID 读卡器被加到 BS 上。由 RFID 和传感器网络提供的无线技术的使用可以在所有类型的家庭和设备中安装健康追踪 2000 系统；RF 波能够穿过墙壁和布通过一个微型发送网络发送重要的信号和位置

信息给中心监测计算机。然后这些信息可以通过因特网很容易地从任何地点获得。这是另一个好用来证明 RFID 和无线传感器网络技术能够在医院外提供普遍的卫生保健作用的例子。

20.3.3 家庭老年人卫生保健的原型

这一节中，我们描述由 Ho 等人[5,11]提出的另一个方案，它建立一个基于集成 RFID 和传感器网络技术的病人护理系统。作者观察到在美国 65 岁以上的人数已经在稳定地增长，在世界的其他地方也是。长寿已经产生了昂贵的与年龄相关的残疾、疾病以及卫生保健。为解决人口老龄化用药需求，作者针对老年患者在家用药提出了一个系统原型。

这个方案包含一个最初的学习阶段和后来的发展阶段。学习阶段通过传感器网络连接一个模拟 RFID 系统来测试技术的兼容性和能力。模拟软件模块描述为提供卓越的学术经历以及在硬件购买成为可能之前是需要的。

在发展阶段，HF RFID（低成本）和 UHF RFID（远距离）的能量同传感器节点一样被利用并应用于一个医疗监测系统。这个系统监测那些老年病人需要的药物的数量并帮助他们服用正确的量。

这个系统包含七个组件——三个节点、一个 HF RFID 读卡器、一个 UHF RFID 读卡器、一个重量秤和一个 BS。下面，会详细描述它们在医疗监测系统中的作用。HF RFID 标签放置在每个药瓶上来识别每个瓶子。HF RFID 读卡器连接跟踪在读卡器范围内的所有药瓶。以一定的时间间隔读取所有的标签，系统能够确定什么时候以及哪个瓶子被移动或者被病人替换。HF 读卡器在这方面的短距离应用还是可以的。重量秤监测秤上药的数量。结合在重量和 HF 标签事件中的变化，当病人吃药时，药瓶和摄取的药量是可以确定的。

下面我们描述追踪病人。一个包含一个读卡器和一个或者多个标签的 UHF RFID 系统用来追踪需要吃药的老年病人。这个病人穿的一个 UHF 标签可以被 3～6m 范围内相关的 RFID 读卡器检测到。子系统选择应用的无线 ID（AWID）UHF RFID 读卡器。病人节点与 UHF 读卡器进行通信来监测到达房间门口或者安装系统的其他地方的病人。因此，这种系统能够确定在附近的病人以及通过警报声警惕病人来摄取需要的药品。

为了提供用户与系统的相互联系，GUI 需要一个显示器。一个嵌入式的显示器可以用于这个目的。由于资源有限，一个基站 PC 软件内的显示器被作为一个替代品。这个显示器模仿并提供一个生动的用户界面（GUI）来帮助病人。这个系统为各种药物/维生素使用大号字体，以及不同的颜色表示药片数量；这将对老年病人更加容易。作为选择，各种药片商标/瓶子的图片可以用来代替药品的名字。

作者已经成功地为监测老年病人在家用药建立了一个原型。HF 和 UHF RFID 技术以及传感器网络会使用。它们指出未来工作将会从药品监测到年老家庭护理系

统，从一个房间到整个房子的原型延伸，使用更多的传感器和RFID元件分布在各种重要的地方以及各种家庭里。另一个扩展是通过邮件增加通知家庭成员的能力以及通过互联网与一个外部健康护理中心监测系统进行联系。

20.4 卫生保健的传感器网络和RFID发展平台

这一节中，提出了集成RFID和传感器网络的旨在卫生保健应用的一种新的应用发展平台。这个工作的初步版本已经在参考文献［17］中提出。在20.4.1节的介绍之后，一种关于编程抽象和相关的中间件方案将会在20.4.2节中进行简要概述。发展平台的核心部分将会在20.4.3节中进行介绍，然后在20.4.4节中介绍协议的执行情况说明，最后在20.4.5节中进行总结。

20.4.1 介绍

目前，大多数RFID和传感器网络应用，包括上面介绍的，执行时是复杂的，低级的方案，指定个别传感器节点的行为[12]。为了使RFID和传感器网络（RSN）实现它们全部的潜能和在商业上取得成功，使RSN用户发展他们自己的应用来适应特定的需求是很有必要的。例如，设计一个照顾病人和老年人的房间，需要卫生保健专家能够配置和控制这些RFID和传感器去感知什么，感知频率、记录的事件以及什么时候发送警告或者采取行动。类似，在追踪医院人员或者血袋中[7]，医院管理人员可以开发一种适合它们具体目的的应用。除此之外，他们可以按照需要改变这些配置和控制参数。

这些RSN用户，诸如上面提到的卫生保健专家和医院管理人员，我们可以称作“领域专家”。他们不是设计RFID、传感器网络的科学家，也不是开发这些传感器复杂程序的软件工程师。他们能够在没有先进科学或者编程背景下开发定制的应用程序将是有益的。

面对这个目标，这节试图提出一种关于RSN的应用发展平台。这个平台由领域专家使用来开发配置和控制有特定目的的RFID和传感器应用。值得高兴的是，这个平台对用户来说是简单的，也有足够的鲁棒性来支持各种复杂的应用。这个工作之前的版本已经在参考文献［17］中提了出来。

设计无线传感器网络中间件的六种方法已经在参考文献［9］中确认（有关WSN中间件的更多讨论会在20.4.2节中给出）。在它们中选择基于事件的方法。它提供了快速的响应和更少的通信开销。这些特性对应用与卫生保健的RSN是有利的。

选择一个基于事件的中间件是很容易使用并能够支持复杂的应用发展的，我们已经选择了基于Java代理的发展框架或者JADE[1]。这是一个用于对等的应用程序的开发和运行执行的基于Java的中间件。它是基于代理模式，可以无缝地工作和

互操作于有线和无线两种环境中。它可以大大简化了需要沟通和协作的自治实体组成的分布式应用的开发。它可以无缝地执行除了Java卡的类型的每一个Java虚拟机（VM）。Java的固有的便携性优势也为RSN增加了实用性，因为它可以简单的迁移Java虚拟机平台之间的应用程序。拟议的框架试图通过尽量减少通信开销和降低计算复杂性来维持一个RSN的平台精神，同时也承认将来这两个RFID和传感器节点的能力将稳步增加。

提出的框架试图维护RSN平台的内在意义，通过最小化的通信代价和简化的计算复杂性，同时掌握RFID和传感器节点的能力在将来会稳步提升。

拟议平台的关键技术贡献包括：①在RSN环境中JADE的应用；②RFID技术，传感器抽象，事件类型；③为应用程序开发的用户界面；④体系结构中使用监测的应用程序状态。

20.4.2　编程抽象及相关中间件项目

这一节中，我们首先讨论RSN中编程抽象和相关中间件项目；然后描述JADE并简单介绍Sun SPOT项目。

20.4.2.1　编程抽象

RSN编程中重要的一方面是编程抽象，即提供传感器和传感器数据的抽象编程。三个主要的编程抽象系统和传感器网络的应用程序已经确定：基于数据的模型，基于代理的模型，宏编程模型[3]。基于数据的模型意味着数据保存在传感器自身中，通过节点之间的检索查询。基于代理的模型与反应过程操作模式是相关的，该反应过程根据传感器的输入和其他反应过程的信息来决定它自己的行为进程。宏编程抽象指的是一个系统编程为一个整体编程的系统带有全局行为，而不是关于点对点的基础。这个抽象是非常有用的，允许应用程序开发人员把系统表示为创造、组合以及状态的转变[21]。

宏编程是所推荐框架采用的方法，它允许不熟悉编程的领域专家使用熟悉的条件即状态来确定系统的行为。除此之外，我们也包括一些由JADE提出的基于代理抽象的方面（参见20.4.2.3节）。

20.4.2.2　中间件

如上所述，作为复杂的、低级程序详细描述个人RFID或者传感器行为的大多数RSN应用软件已经实现。显而易见，一个单一的硬件平台将很可能满足不了广泛的可能性应用。为了避免应用程序特定硬件的发展，因此最好用一（小）套带有不同功能的平台来覆盖设计空间[4,12]。

中间件是隐藏分布式系统复杂性的标准做法。在RSN中，中间件从RFID标签和传感器收集大量数据，并以适当的格式来为目标分布式应用程序进行数据汇总和管理[4,21]。

Yoneki和Bacon已经对传感器网络中间件做了一个广泛的调查[21]。他们确定

了六种方法：数据驱动、基于事件、面向质量服务（QoS）、面向因特网、基于代理和集中的；他们也总结了使用各种方法的系统。在这些方法中，已经把基于事件的方法选择为我们推荐的框架。它可以作为为 RSN 广泛提供出版/订购服务的系统，这里订购者是应用程序，出版商是 RSN 节点。事件发生用来响应监测现象的变化，并且随后报告给订购者。这种方法积累少量通信开销；一个单独的订阅信息发送给 RSN 能够为应用的周期提供事件信息。在 RSN 中，当通信成本都处于溢价时，这种方法是非常可取的。其他采用基于事件的方法的中间件包括 DSWare[8]、Mires[13] 和 Scope[14]。

20.4.2.3 JADE

JADE 是一个基于 Java 使用代理模式的中间件。代理是一个自主的、主动的软件组件，在分布式环境中进行对等的编程活动。基于事件的系统是用来与代理互联的中间件平台[1]。JADE 也是智能物理代理基金会（FIPA）相容代理平台，允许与其他兼容平台的 RIPA 协同工作。JADE 还提供了独立于基础网络和 Java 版本的应用程序编程接口（API），从而允许设计的统一和可移植。

JADE 包括如白页和黄页服务的 FIPA 指定服务。白页服务是全球所有代理的名称和地址列表。黄页服务又称 JADE 目录服务商（DF），包含该代理商提供的服务信息。代理注册服务伴随着描述服务的语义信息。需要服务的代理可能通过 DF 使用语义查询来查找代理平台内的有效服务[2]。

JADE 中的代理功能通过行为进行编程（注意，“行为”非“行为”是使用 JADE 环境的正确拼写）。每个 JADE 代理执行自己的线程，然后按照顺序执行程序员提供的行为。JADE 提供了一个为具体的行为延伸的可扩展基类行为。

如 20.4.1 节所述，最好有一个 RSN 的平台，不仅使用简单，而且能够支持复杂应用的发展。JADE 能够在大多数 Java VM 上准确地执行，包括 J2EE、J2SE 和 J2ME[1]。此外，它可能被视为轻量级的中间件：一个微型版本的 JADE、LEAP，有与 JADE 相同的 API 并且在 J2ME 上运行（资源约束计算环境的 VM）。它兼容连接的有限移动设备（CLDC）/移动信息设备概况（MIDP）1.0 环境，常见于资源受限的设备（如移动电话和掌上电脑）。JADE 在 MIDP 1.0 环境中运行时内存大约为 100KB，但使用“ROMizing”技术（即 JADE 与 JVM 编译一起）可以进一步降低到 50KB[1]。此外，JADE 提供多余的白页和黄页服务，以及在出现故障时切换和发现的方法。这增加了该系统的鲁棒性，这也是 RSN 应用程序所需的另外一种特性。

Sun 实验室最近开发的 Sun SPOT（小型可编程对象技术）硬件系统，专门用于支持无线传感器和传感器（执行器机制相结合传感器）的应用发展[15]。它支持 Java 和通用的集成开发环境（IDE），并具有 Squawk VM，一个小的 J2ME VM。它还允许“在金属上”运行这些应用程序（直接在中央处理单元上，而不使用基础操作系统（OS））。这是一个很适合推荐的工作的硬件平台的例子。

20.4.3 应用程序开发平台

推荐的框架是一种发布/订阅系统建立在JADE平台提供的服务所发现的框架之上。RFID和传感器是JADE平台运行的一个实例，使用JADE黄页服务登记它们的存在。应用程序可能再订购它们的服务。

这一节首先叙述了准备工作和数据结构，然后详细介绍一个应用程序的开发过程，并就能量管理的一些评论进行了总结。

20.4.3.1 准备工作和数据结构

20.4.3.1.1 传感器登记和服务发现

在该平台的第一步是服务发现。所有的RFID和传感器必须登记它们的服务。RFID代理人或传感器代理在JADE黄页服务为它们的服务做广告和等待订阅。登记时，它提供的信息将对应用程序开发者有用。

RSN可能由大量的传感器补充，它们在感应能力和复杂性中变化。我们已经确定了三个作为更普遍的抽象的传感器分类：基于定量的、基于状态的和基于身份的（将在20.4.3.2节中详细讨论）。表20-1列出了传感器在服务登记时提供的信息。注意，一个RFID/传感器可以提供多个类别相关的服务。例如，一个RFID传感器很可能除了基于识别的服务（识别被检测的对象）提供一个数量上的服务（报告被测对象的人口规模）。

表20-1 RFID和传感器在服务广告期间的信息报告

	RFID/传感器分类	RFID/传感器类型	位置	测量单元	测量精度	样本率	状态调整
基于定量	×	×	×	×	×	×	
基于状态	×	×	×			×	×
基于身份	×	×	×			×	

20.4.3.1.2 数据模式

每个RFID或者传感器维持自己的感知数据微型数据库。该数据库由数据样本和时间戳构成。在表中存储的数据量不同于目前使用的RFID和传感器应用的要求。RFID或者传感器也可维持其功能的描述，包括表20-1中描述的数据。

20.4.3.1.3 事件类型

三种RFID/传感器各种有一套相关的事件。每种传感器的例子事件在表20-2中所示。

除此之外，表20-3显示了RFID读卡器一个事件的例子。通过识别相关传感器事件类型，我们能够做一个可扩展的系统，这里一整套的传感器类型不需要在设计的开始阶段被确定。表20-2中的每种事件类型通过系统中的一个行为（JADE中的行为指的是20.4.4.1节和表20-1）所描绘。在具体抽象事件和行为之间有一个一

对一的映射。

表 20-2 传感器事件类型的最小设置

传感器类型	事件类型	输入到事件发生器	数据绑定到事件	描述
基于定量	门限	门限值	数据值，时间戳	传感器值增加/减少超过门限
	滑动窗口门限	门限、窗口大小	传感器值，时间戳	传感器值增加/减少大于窗口门限
	平均窗口门限	门限窗口大小	窗口平均后，时间戳	窗口平均值增加/减少超过门限
基于状态	输入状态	状态 S	当前状态，时间戳	传感器已经输入 S
	退出状态	状态 S	当前状态，时间戳	传感器已经退出状态 S
	状态改变	状态 S，状态 T	当前状态，时间戳	传感器从状态 S 转变为状态 T
	状态门限中的时间	状态 S，时间 t	当前状态，输入时间时状态，时间戳	传感器处于状态 S 的时间超过 t
	提前退出	状态 S，时间 t	状态 S，新状态，时间戳	传感器在超过时间 t 前退出状态 S
基于身份	对象检测	对象识别	对象识别，时间戳	先前识别对象检测
	新对象检测	N/A	对象识别，时间戳	新对象检测
	对象去除	对象识别	对象识别，最后检测时间戳	最后一次传感器观察检测的对象不再会被检测
	时间参考对象去除	对象识别，超时间隔	对象识别，最后检测时间戳	先前传感器观察中被检测的对象在比超时间隔长的时间不会被检测

表 20-3 RFID 读卡器的事件设置

	事件类型	输入到事件发生器	数据绑定到事件	描述
基于身份的传感器事件	RFID 标签检测	标签 ID	标签 ID，时间戳	先前识别的 RFID 检测
	新的 RFID 标签检测	N/A	标签 ID，时间戳	新的 RFID 标签检测
	RFID 标签去除	标签 ID	标签 ID，RFID 标签最后检测的时间戳	最后一次 RFID 阅读器轮询中检测到的 RFID 标签不会再被检测
	RFID 标签去除	标签 ID、超时间隔	标签，RFID 标签最后检测的时间戳	先前轮询中已经被检测到的 RFID 标签在多于超时间隔的时间内不会被检测到

（续）

	事件类型	输入到事件发生器	数据绑定到事件	描述
基于定量的传感器事件	数量大小门限	数量门限	当前数量大小，时间戳	数量大小增加/减少超过门限
	滑动窗口数量大小门限	数量门限，窗口大小	当前数量大小，时间戳	数量大小增加/减少超过窗口的门限
	平均窗口数量大小门限	数量门限，窗口大小	窗口的平均数量大小，时间戳	窗口平均数量大小增加/减少超过门限

20.4.3.1.4　基于事件的发布/订购系统

黄页的界面和发布/订阅系统的提出是基于三种类型的传感器。一定数量的 RFID/传感器产生并发布测试一些环境的数字数据，就像温度传感器报告摄氏度的温度。另外，温度传感器可以提供基于传感器的条件解释的状态数据，过热或过冷。同样，一个基于识别的传感器如 RFID 读卡器，返回一个感应对象的确认信息。

20.4.3.2　应用发展进程

系统的主要目标之一是去除 RSN 应用软件高水平发展的编程语言，允许领域专家在不熟练编程语言的情况下可以直接使用 RSN 软件。这个目标的基础是扩展由 JADE 提供给 RSN 领域的服务发现系统。在基础的上面，我们建议需要一个描述应用软件的事件成分界面，按照下面的步骤。注意，这种进程不仅减缓了应用软件的发展，也使得应用软件在不太困难的情况下被改变和扩展。

Step 1：创建应用程序

开始一个新的应用软件，开发者（领域专家）首先要为这个应用软件创建设计空间。它将会提供一个区域来存储选定的 RFID 和传感器、事件订阅、分组活动和选择执行器以及事件执行器关系的联系，如下所述。

Step 2：传感器的选择

应用软件的开发者首先给系统提交一个查询，要求他们感兴趣的传感器类型。更多复杂的问题涉及对位置，采样率等的限制。把一套满足问题所描述的系统规定参数的 RFID 和传感器提供给了开发者，然后再选择有用的传感器。

Step 3：事件订购

回想每种类型的 RFID/传感器有一套相应的抽象事件（见表 20-2）。开发者然后浏览并选择有用的事件。事件的选择过程将被简单地标记一个复选框，给事件输入必要的信息生成功能，并点击订阅按钮。然后，当应用程序启动时，期望事件的列表存储订购申请。

Step 4：有限状态机器的事件组

一旦感兴趣的事件已经被确认，它们就能够分类形成一个综合事件。使用事件来触发转变，开发者能够通过一个有限状态机（FSM）描述一个综合事件，代表复

杂事件的相互关系，到达最终的状态导致一个新事件的产生[6,8]。

对于综合事件的产生，Li等人描述绝对有效间隔作为一种方法来满足事件对时空关系的需求[8]。我们已经采用了这种方法。当定义描述综合事件产生路径的FSM时，应用程序的开发者能够把休息间隔放在每种状态上，这也能够触发转换。通过这种方式，有关时间关系的事件限制能够精确地表达。然后，FSM能够被一个由JADE专门提供执行类似FSM的行动的行为所描绘。

Step 5：执行器的选择

类似于RFID和传感器的选择，开发者提交一个满足一系列规定参数的请求执行器的问题。系统规格参数包括执行器型号，位置等。然后，开发者选择希望在他的应用中使用的执行器。

Step 6：事件执行器关系

为了触发一个行为，应用程序开发者把一个事件连接到一个执行器上并设置执行器的执行参数。执行器对请求采取行动；即当连接到执行器的事件产生时，一个行为请求会被发送到执行器。依靠指定的行为，结果可能是一次行为或者由于另一个事件的终止而产生的连续行为。

20.4.3.3 能量管理

众所周知，能量管理在大多数传感器网络应用中是很重要的。在所提出的平台中，只要可能，能量消耗都尽量最小化。接下来我们讨论一下关于这个的两种方法。

20.4.3.3.1 本地计算最小化通信

众所周知，在能量消耗方面，通信要比计算消耗得多。因为这是一个事件驱动系统，无线通信需要订购传感器事件，然后当条件满足时，通知观察者有关事件出现的情况。通过在每个设备上（指的是表20-1和20.4.2节）创建一个本地设备代理（DeviceAgent），推荐的平台保持大部分本地计算，最大限度地减少传感器采样和无线通信，因此来减少能量消耗。

20.4.3.3.2 控制传感器采用和报告率

在发布/订购模式中，维持一个最小的消息传递开销是至关重要的。一个手段是控制采样率。因为每个传感器有一个决定最大采样率的最小采样间隔，所以仔细控制这些最小采样间隔是很重要的。很多应用程序不需要传感器在它最大的速率下采样和报告。允许采样率指定于订购消息中，有效地控制了传感器数据通信，因此减少了能量消耗。

20.4.4 原型实现

我们已经开发出一种简单的，作为一个概念证明功能样机。原型是由三个模块组成。第一个模块是一个基于代理的传感器事件登记和JADE上的环境监测。第二个是在第一个模块的顶部制定一个发展应用的GUI。第三个是要有可观察的现象、

设备和传感器可以安排形成测试环境的实验环境。先对这三个模块进行了描述，接着针对在原型中实现应用程序进行了讨论。

20.4.4.1　核心模块：登记和监测

平台的主要作用部分是基于代理的 RFID/传感器事件登记和环境监测；一种图的核心组件如图 20-1 所示。一切都集中在 DeviceAgent 类中，并在 DeviceAgent 和设备之间有一对一的关系。每个 DeviceAgent 反过来连接到一个 PlatformAdapter 上，这是一个一般的界面，通过使用的各自硬件平台来实现。在这个原型中，我们模仿了底层硬件；扩展计划包括实际硬件。通过确认标准界面，新的硬件平台可能在不改变核心模块的情况下增加。

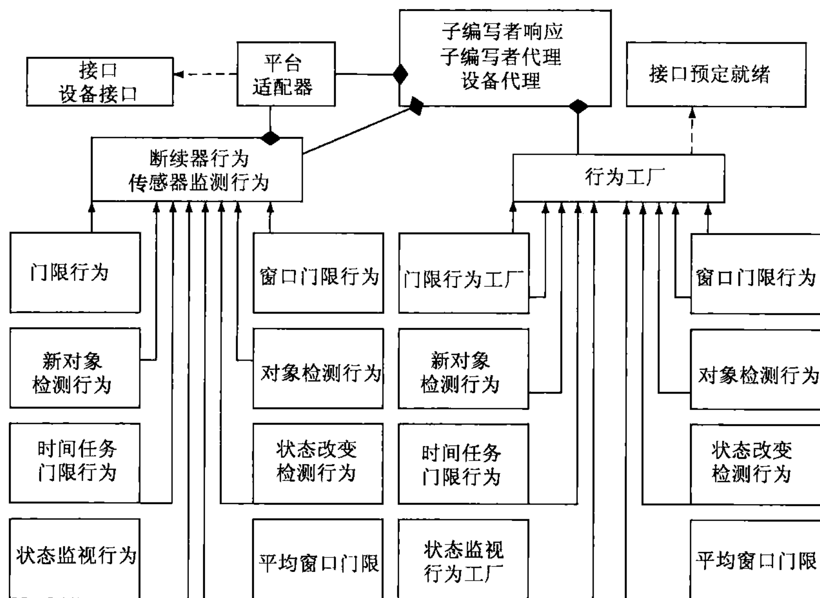


图 20-1　传感器监测模块的核心类图

抽象的 PlatformAdapter 类为设备的相互作用提供了许多方法。为了保持界面简单又灵活，一个 SensorProfile 类被创建来囊括所有依附于设备的 RFID/传感器的详细资料。创建时，一个设备代理将调用 PlatformAdapter 的 getSensorProfiles（）方法，并且将会为每个 RFID 或者传感器收到一系列的 SensorProfiles。PlatformAdapter 也提供一种方法来设置传感器界面，允许每个 RFID 或者传感器的特性随着时间改变。除此之外，它给 RFID 或者传感器提供数据的两种方法被设备代理（DeviceAgent）感应监测行为（SensorMonitorBehaviors）所检索：只读最近收到的数据或者检索所有及时数据，返回作为指向一系列 SensorData 的迭代器。注意，SensorData 是另一个抽象类，它的基类分为数量数据（QuantityData）、状态数据（StateData）和识别数据（IdentityData）。

一旦被设备代理收到，SensorProfile 的特性就会用由关联设备的代理提供作为服务的 DF（黄页服务）登记。DF 提供有效地方法来搜索和选择对应用程序开发者感兴趣的设备。开发者能够查询 DF 的带有一种特别传感器的设备，然后用设备代理登记来监测由传感器产生的数据。每个设备代理执行 JADE 定义的认购管理界面，反过来有许多认购响应行为运行在它的线程。

参考表 20-1，在运行于代理上的订购响应（SubscriptionResponder）行为和代理知道的行为工厂（BehaviorFactories）之间有一个一对一的关系。行为工厂用来产生行为，这些行为实际上监测传感器数据和返回订购者感兴趣的传感器事件。为了开始一个行为，应用程序的开发者发送一个订购信息给感兴趣的设备代理。订购信息里包括目标传感器的传感器 ID 和一些特别是事件产生行为感兴趣的其他数据条目，如采样率、阈值、窗口大小、状态名称和对象 ID，其中事件 ID 把订购信息连接给订购响应行为（以及随后给正确的行为工厂）。应用程序的开发者能够决定什么类型的信息是每个事件发生器所需要的，通过检查 DF 中提供的信息。

20.4.4.2 图形用户界面（GUI）应用程序开发

应用程序可以通过利用 JADE 平台提供的 FSM 行为很容易的进行开发。一个应用程序只不过是带有进入新状态的各种关联行为的一套状态（行为可以在将来的扩展中赋予状态出口）。回想 JADE 有一个模仿 FSM 行为的行为子集。每个 JADE FSM 状态包含另一个功能的 JADE 行为（可能甚至是另一个相当复杂应用程序的 FSM 行为）。状态转换通过返回值被触发，在它生命周期结束的时候产生一个行为。每个应用程序的状态是状态监测行为或者一些多数行为的结合。那些触发转换的返回值是事件 ID，它是在或者事件条件满足、超时出现、或者检测到失败时通过这些行为返回的。

我们创建了一个简单的面向例子程序（将在下节中介绍）的具体应用——GUI。GUI 是为了选择 RFID 和传感器，改变与之前规定的具体应用 FSM 相关的特性。注意，一般的应用程序开发 GUI 也是可能的，并且为将来留有扩展。

20.4.4.3 实验环境

实验环境已经被建立，包括可观察的现象、RFID、传感器和其他无线设备。由于 RFID 和传感器服务分成了三类，因此也创建了三种现象：基于数量的、基于状态的和基于识别的。

环境通过简单的 GUI 管理，它管理的环境包括可观察的模拟设备、可运行的 RFID 和传感器以及可运行可观察的现象。然后，每个设备与通过 PlatformAdapter 和 DeviceAgent 相互作用的 DeviceAdapter 交互，如 20.4.4.1 节所述。

20.4.4.4 应用例子

为了进一步说明推荐的开发平台，这一节中，我们将要描述一个把平台应用于卫生保健设备的例子。它基于之前实现的硬件原型，如 20.3.3 节所描述的，老年病人的医疗帮助系统。推荐的框架可能在例证中显示是有用的。

考虑到一位老年病人的家里，可以装备传感器来监测病人的健康和行为。RFID读卡器放置在病人家里的不同门口来监测环境。带有重量传感器的专门刻度代表了病人目前的体重。遍及家里的声音传感器用来检测病人的移动、活动或者呼叫求助。冰箱里的光传感器用来推断病人吃饭的频率。

使用我们的系统，卫生保健的提供者能够开发一种应用软件来监测病人的行踪，根据检测的重要事件来定义行为。一个每次通过RFID读卡器传递病人去睡觉还是起床的基于识别的事件，用来跟踪睡觉的样式以及当预期的监测持续太久时发送潜在的疾病信号。刻度用来检测任何剧烈的重量变化，能够发送潜在疾病的信号。冰箱光事件之间一个过度长的间隔将触发一个病人不在吃饭的事件。一个突然尖叫的声音触发一个可能的跌倒或者呼叫求助的事件。根据产生的事件，然后应用程序采取行动，包括呼叫卫生保健者不健康的变化、打电话或者给邻居或者健康检查的亲属发封邮件，以及在极端的条件下请求紧急的医疗援助。

20.4.5 摘要

这一节中，我们已经论证了有关卫生保健的RFID和传感器网络的应用软件开发平台，建立在JADE中间件的顶层。它提供了一套简单的、一般的独立应用的RFID/传感器和事件，卫生保健人员可以轻松地组建和提高简单高效的应用程序。未来工作可能包括通过使用一个类似于Java本地界面（JNI）的界面把平台扩展到其他无Java的平台，以及应用程序的开发和维护的一般的高水平GUI的开发。为了驱动的实现，一个类似于PlatformAdapter的界面可以开发来增加硬件和软件驱动能力。最后，一个数据驱动设备可能要增加来补充事件驱动系统。

20.5 结论

RFID和传感器网络都是令人兴奋的领域；集成两种技术在不远的将来会提供更多的潜力。这一章描述了一些使用结合技术的卫生保健应用。随着他们的研究和开发成就不断增强，以及全球的老龄化人口稳定的增长，我们期望看到更多令人兴奋的应用这两种以及其他正在兴起的卫生医疗和医疗科学其他领域的无线技术的建议。

参考文献

1. Bellifemine, F., Caire, G., Poggi, A., Rimassa, G., JADE A While Paper, http://exp.telecomitalialab.com, September 2003.
2. Bellifemine, F., Caire, G., Trucco, T., Rimassa, G., JADE Programmer's Guide version 3.3, http://jade.tilab.com/doc/programmersguide.pdf, March 2005.
3. Bonnet, P., Gehrke, J., Seshadri, P., Querying the physical world, *IEEE Personal Communications*, 7(5), 10–15, October 2000.

4. Heinzelman, W.B., Murphy, A.L., Carvalho, H.S., Perillo, M.A., Middleware to support sensor network applications, *IEEE Networks Magazine*, 18(1), 6–14, January/February 2004.
5. Ho, L., Moh, M., Walker, Z., Hamada, T., Su, C.F., A prototype on RFID and sensor networks for elder healthcare: Progress report, *Proceeding of the 2005 ACM SIGCOMM Workshop on Experimental Approaches to Wireless Network Design and Analysis*, Philadelphia, PA, 2005.
6. Kaston, O., Romer, K., Beyond event handlers: Programming wireless sensors with attributed state machines, *Fourth International Symposium on Information Processing in Sensor Networks* (*IPSN 2005*), Los Angeles, CA, pp. 45–52, April 2005.
7. Kim, D.-S., Yoo, S.K., Kim, H.O., Chang, B.C., Bae, H.S., Kim, S.J., Location based blood bag management using active RFID and ubiquitous sensor network, *6th International Special Topic Conference on ITAB*, Tokyo, pp. 320–322, 2007.
8. Li, S., Lin, Y., Son, S., Stankovic, J., Wei, Y., Event detection services using data service middleware in distributed sensor networks, *Information Processing in Sensor Networks Workshop*, Pao Alto, CA, 2003.
9. Madden, S., Franklin, M.J., Hellerstein, J.M, Hong, W., The design of an acquisitional query processor for sensor networks, *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, San Diego, CA, pp. 491–502, 2003.

第21章　应用于建筑物结构监测的RFID与传感器网络的集成

自动结构“健康”监测（SHM）试图通过使用一套嵌入式传感器持续地评估建筑结构的完整性。这一章中，我们讨论电阻应变计在结构健康监测中的应用，这里应变计用来测量关键结构部件在一个或者多个尺寸中的变化。我们还提出一个二进制应变计用来检测明确超过预定阈值宽度的裂缝。二进制应变计可用在没有复杂数据分析和相关数据报告、结构分析结果或者结构图的信息的情况下描绘直接损失。在网络要求方面、二进制应变计数据报有一个低速率和自然弹性的数据丢失。我们讨论商用的现成（COTS）数据采集和通信设备如何用于建立一个多跳的无线结构健康监测网络。SHM网络具有成本低、可配置以及适合灵活的部署的特点。

21.1　概述

如果世界贸易中心大楼装备了温度传感器，“9·11”事件的灾难就可以部分地避免。工程界似乎已经对大楼的坍塌达成了一致，它是由于高温引起承载能力下降所导致的。如果工程师知道建筑里的温度就可以协调营救工作，派遣人员的决定可能就会转变。随着传感和通信技术的发展，一系列监测建筑结构的传感器持续主动地成为可行的。在地震、风暴或者爆炸等极端事件期间，传感器阵列报告的损失信息将会及时通知营救/响应操作和利益相关者。没有这些阵列，损失检测必须得通过亲眼检查。这个过程既费时又昂贵，因为它需要有经验的人员。

自动结构“健康”监测（SHM）的目标是使用一套嵌入式传感器持续地评估结构的完整性，并把结果实时地发送到远处的控制中心。控制中心分析数据报告来检测传感器信号的变化，指出损坏或其他安全问题。与结构健康相关的物理特性包括张力、压力加速度（涉及结构的振动）和裂缝宽度。基于振动测量的SHM网络包括［KPC07，XRC04］，测量通过利益结构上安装的加速计。这些网络通过检测结构硬度的变化可以推断出危险性。有关保证结构安全的硬度损害并不是没有价值的，因为当结构受到大的负载时，它们就会变形。遭受硬度上改变的结构，可能会根据设计相应地有影响。另一方面，一个大的结构上的部分个人局部的失败可能在整体结构的硬度上不会产生大的变化，并且可能因此避开检测。就网络需求而言，基于振动测量的SHM系统需要不同传感器报告的数据间紧迫的时间同步。高分辨率加速度计的数据需要以100Hz或更快的速率进行采样。因此所要求的网络应该有一个高的数据率。

这一章中，我们讨论SHM中电阻应变计的不同使用。应变计用来测量关键性结构方面一个或多个尺寸的变化。我们还提出一个二进制应变计用来检测明确超过预设阈值宽度裂缝。这些应变计可以在数据报告、结构分析结果或者结构图信息之间没有复杂的数据分析和相互关系的情况下直接描绘损害。就网络需求而言，二进制应变计数据报告数据率低，并且对数据流失有固有的弹性。我们讨论COTS数据获得和通信设备怎么样用来建立一个多跳的无线SHM网络。SHM成本低、可配置并且服从灵活、易用的部署。

21.2　电阻基传感器背景

电阻基传感器通过联系尺寸和电阻的变化来工作。导体的电阻是由横截面面积A、长度L和电阻率ρ决定，即

$$R=\frac{\rho L}{A} \tag{21-1}$$

如果导体在轴向上受力，那么长度L跟施加的压力成正比。横截面积A也随着轴向压力改变。如果受到的是张力，面积减小，如果受到压力面积增加。对于小的压力，面积的相对改变直接与材料的泊松比成正比。长度和面积的相对改变都会使电阻变化。压力也会影响材料的电阻率ρ。电阻率随着张力增加，随着压力减小。

对于小的压力（小于8%［DAL93］），与压力相联系的电阻总的变化ΔR可以表示为

$$\Delta R=S\varepsilon R \tag{21-2}$$

式中，S是一个比例常数。对于某些镍基合金（康铜、噶玛、镍铬等弹性的）和其他合金（Armour D和铂钨），S从2~4变化［DAL93］。这些合金电阻率高，范围广，其中电阻和压力的关系仍然是线性的。

21.3　电阻应变计

式（21-2）表明与给定的压力相关的电阻随着导体和电阻R的增大而增大。为了增加长度和减轻处理与使用，导体被照片蚀刻到薄底板材料的格子上，如图21-1所示。

图21-1中所示的设备叫做电阻应变计或者简单应变计。生产应变计的许多配置具有不同几何形状、合金和背景材料。应变计被仔细地依附在测量应变力的介质上。黏合剂、背景材料、合金配置和应变计的几何形状是影响应变计灵敏度的所有因素，例如应变的相对电阻率变化。因此，电阻应变相对变化之间的比例常数与S不同：

$$\Delta R=S_G\varepsilon R \tag{21-3}$$

市售应变计的有效灵敏度 S_G 大约是 2.0[MG07a]。

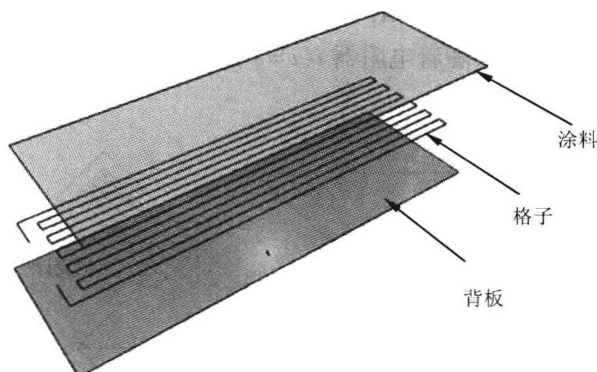


图 21-1　应变计的组成

由于半导体在应变计中的配置，有限长度方向与张力的方向垂直。这使应变计对横向张力 [MG07b] 很敏感。市售的有效的应变计被大约等于 -0.3 乘以轴向力的横向力的领域所校准。应变力的测量使在不同的条件下，需要更改应变因素 S_G。对于一般的测量仪，忽略横向灵敏度导致的错误对于 -1 ~0.5 之间的横向对轴向应变率不可能超出 5%。应变计做的测量也对温度变化、用来连接应变计到信号控制系统的线上电阻的变化、噪声以及湿度（可能影响黏合剂的稳定性）很敏感。

设计应变计来测量由机械动作（力和力矩）引起的本地应变，通常不超过 10%。如果用来直接测量压力，应变计会被附在正在监测的结构上。应变计也用在各种不同的应用中，它们也是传感器的一部分，转化应用于结构的机械动作为电信号。这些应用中，应变计被附在一个在机械动作下会变形的元件上被监测。使用应变计的传感器包括测压元件和压力传感器。应变计还用于变形测定器测量有限长度的变形。

21.4　信号调节电阻应变计

在一般使用应变计的应用中，通常更青睐于测量电压的变化而不是电阻的变化。一个叫做惠斯顿电桥（Wheatstone Bridge）的电路用来把电阻的变化转化为电压的变化。这个电路如图 21-2 所示。它由串联起来的四个电阻（R_1、R_2、R_3、R_4）组成一个环。两个电阻之间的连接叫做一个节点。稳定的电压用来在两个相对的桥节点之间建立一个潜在的不同 V_S。横跨对应桥节点的剩下部分的潜在差异是 V_o。如果 R_1R_3 的乘积等于 R_2R_4 的积，那么 V_o 就等于 0。

由于应变或者其他因素导致惠斯顿电桥的电阻器变化时，V_o 的输出将会偏离 0。电桥中的电阻器的阻值变化引起的电压变化是

$$\Delta V_o = \frac{r}{(1+r)^2}\left(\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4}\right)(1-\eta)V_s \tag{21-4}$$

式中，r 是 R_1/R_2 的比；R_i 随着电阻器 $i(i=1,2,3$ 和 $4)$ 相关的电阻变化；V_s 是激励电压；η 是非线性因素。

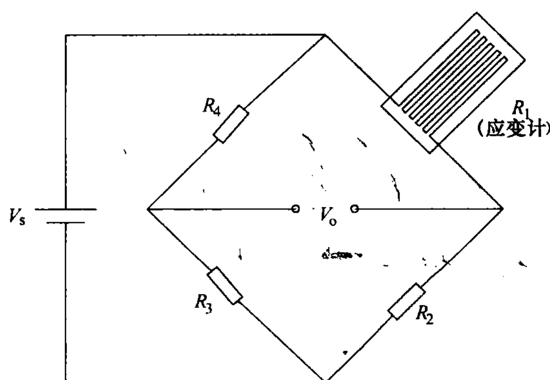


图 21-2 惠斯顿电桥

对于大多数应用，η 不会超过应用压力的 1.25 倍。单电阻器应变桥中，η 大约等于施加的压力。

注意，如果所有的电阻器的电阻变化相同，输出等于零，因为与相关的电阻变化的相邻桥的信号是相反的。事实是可以用来滤除噪声和与监测的机械行为不相关联的电阻的变化。但是在我们详细讨论这之前，我们要考虑惠斯顿电桥产生的重要信号的顺序。土木工程中，应变很少小于 0.0001。当 $r=1$ 时，单一的有源应变计（“四分之一桥梁”的配置）0.0001 的应变输出电压的顺序是

$$\Delta V_0 = \frac{1}{4}(2\times 0.0001)V_s = 0.00005V_s \tag{21-5}$$

当激励电压高达 10V 时，我们需要准确地测量 0.0005V 的低电压。理想情况下，需要测量电压的设备分辨率应该第一百个期望的测量值：5×10^{-6}V。我们的结论是需要放大电桥产生的信号，以及使用稳定、精确的设备和电阻器。精确稳定的监测设备和电路方面的需求对移动、结构紧凑以及易于部署应变计网络的实施构成了挑战。其他涉及传统应变计技术使用的挑战如下：

① 温度灵敏度；

② 湿度灵敏度；

③ 噪声；

④ 布线；

⑤ 开裂。

系统对温度的灵敏度与①应变计本身的灵敏度变化，②电阻的变化以及③结构

温度与应变计感应的体积变化差异有关。如果结构热量膨胀的因素与应变计的不同，那么这些差异就会发生。可以选择应变计来使热量膨胀系数的差异变小[MG07c]。温度灵敏度也可以通过使用图21-3所示的惠斯顿电桥的配置来减小。

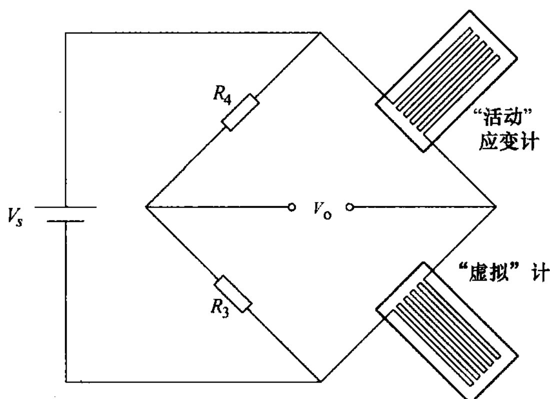


图21-3　带有"温度补偿"或者"虚拟物体"的惠斯顿电桥

在显示的配置中，两个电阻器受到应变力，两个没有。没有受到应变力的两个电阻器形成了信号调节设备的一部分。其他两个电阻器是应变计。一个依附于结构并且对结构中的压力进行回应。另一个应变计（通常称为"虚拟"计）依附在由相同材料制作的原件上，这个结构不受压力的作用。因为惠斯顿电桥中邻近的电阻器信号会互相抵消，如果它们没有引起结构中的压力，温度的变化将不会产生输出。如果贴在结构上的应变计包含由于温度变化引起的压力和结构中的相关压力，那么与温度相关的信号原件会被贴在不受压力的原件邻近信号取消。

在许多应用中，感应计位于远离信号调节系统的地方。感应计和信号调节系统通过引线连接。温度的改变可以导致这些线的电阻的变化。这些电阻的变化可以产生假的输出。但是"温度弥补"或者"虚拟"计使用背后的相同概念可以用来减少引线电阻变化的影响。这样做，如果单个计是可以的，图21-4中所示的配置会用来把它和信号调节电路连接起来。

在所示的配置中，电线1和2电阻的相对变化会相互抵消。如果用来测量输出的仪器有高的阻抗，那么电线3电阻的变化不会影响输出。

引线产生了另一个挑战，因为它们影响系统的灵敏度。它们不会对结构的压力产生影响。但是它们影响$\Delta R/R$［式（21-4）］中的分母。本质上讲，R不等于应变计的电阻，而等于加线电阻的应变计电阻，$R_{线1}$：

$$\frac{\Delta R}{R}=\frac{\Delta R}{R_{计}+R_{线1}} \tag{21-6}$$

电阻的变化依然是

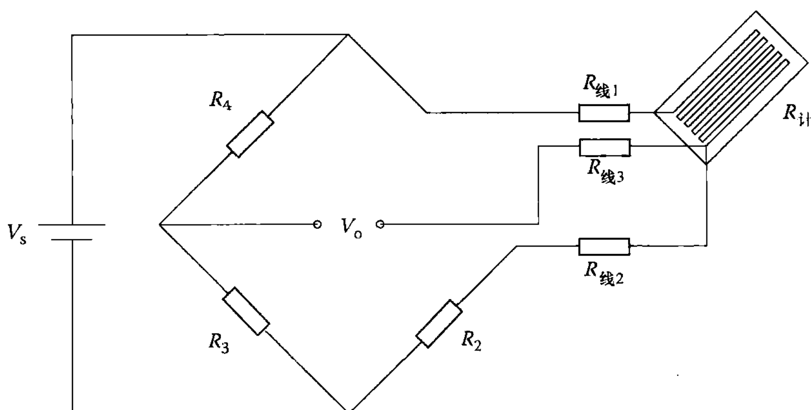


图 21-4　三个引线连接

$$\Delta R = S_g \varepsilon R_{计} \tag{21-7}$$

从式（21-6）和式（21-7）中，整理得

$$\frac{\Delta R}{R} = S_g \varepsilon (1 - L) \tag{21-8}$$

这里

$$L = \frac{R_{线1}}{R_{计} + R_{线1}} \tag{21-9}$$

L 表示灵敏度的“流失”。如果引线的电阻已知，系统的输出可以通过使用邻近等于 S_g（$1-L$）的因素校准。如果引线的电阻是未知的，系统可以通过连接类似于 $R_{计}$ 或 R_2（见图 21-5）的电阻器来校准。这个电阻器的连接导致了阻值 ΔR 的变化，类似于应变导致的电阻变化。如果电阻器选择来产生一个已知的阻值变化，然后系统的灵敏度可以被调整使输出应变等于相关感应电阻的应变为 $\Delta R/(R_{计} S_g)$。这个步骤叫做“分流”校准。

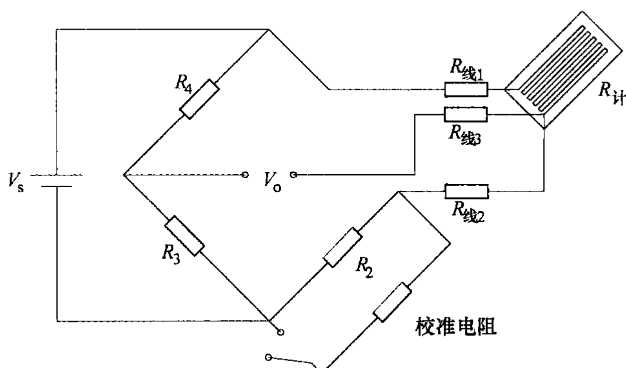


图 21-5　分路校正

21.5　大应变二进制输出电阻基传感器

惠斯顿电桥用途很广，可以用在许多依赖感应用途的配置中。如讨论所述，它有局限性，但许多可以通过仔细地配置或者校准进行处理。但是结构的开裂代表了一个关键性的挑战。横穿应变计的裂缝会阻止应变计，并且惠斯顿电桥部分是无用的。惠斯顿电桥电路的分裂会在它的输出产生大的变化。这是一个弊端，因为导致裂缝变形的重要性是不能被估计的。如果关心的是只对裂缝的检测，这也是一个优势，因为与形成裂缝相联系的输出的变化是很容易被识别的。这个想法促进了一种不同型号计的发展。这些应变计唯一的目标是检测裂缝而不是估计应变。如果唯一的目的是检测裂缝，那么系统精度的要求可以放宽和数据采集过程可以简化。

Wood 和 Neikirk[WN01]使用 EAS 标签，Morita 和 Noguchi[MN06]使用 RFID 标签进行裂缝检测。两个系统基于简单的想法，如果在元件中形成裂缝，依附于结构元件表面的导体可以被折断。这些导体在为 EAS 标签或者 RFID 标签供能的自感应电路中用作开关。切换开关可以使 RFID 标签不能与外部读卡器通信或者在由 Wood 设计的 EAS 标签的情况下，它可以改变电路的共振频率。每个传感器产生的信息是二进制的，并且指示导体是否断裂。这些技术已经被证明对于检测结构元件相对狭窄的裂缝（厚度不超过 2mm）是有效的。有一些应用不需要甚至不想要检测小厚度的裂缝。特别是在钢筋混凝土结构上，狭窄的裂缝是不可避免的，并不表明结构的完整性已经受到影响。钢筋混凝土的相关性是检测宽的裂缝和大的中断(见图 21-6)。



图 21-6　对钢筋混凝土柱的破坏

Chin 等人［CRM08］开发了一种传感器，在结构性元件一维度方面引发了一场大的变化。传感器如图 21-7 所示，称为“CRM 计”。CRM 计可以用来检测厚度超过预期门槛的裂缝。这个门槛可以依靠可承受损害的标准来选择。随着长度超过 12in，CRM 计可以覆盖一个足够大的面积给予结构元件状态可靠的评估。

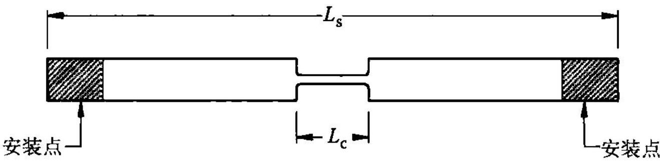


图 21-7　CRM 计

CRM 计包含带有在窄薄的“颈”的可传导层压材料。这个颈可以用不同的材料或者剩余层压制件的相同材料进行人工制造。应变计的总长度 L_s 决定多大的一个区域可以被单个应变计监测。颈的长度指的是计的长度 L_c。应变计连接到其两端使用标准技术增加电阻应变计的结构元件上。两个连接物的变形主要集中在颈部。这些变形与结构元件的应变和裂缝有关。选择应变计的长度 L_c，如果连接点间的总变形超过预设的值，应变计就会折断。实验表明导致颈部折断的变形与 L_c 有关。应变计可以用作开关启用或者禁用无源 RFID 标签（见图 21-9）或者感应电路的开关（见图 21-8）。

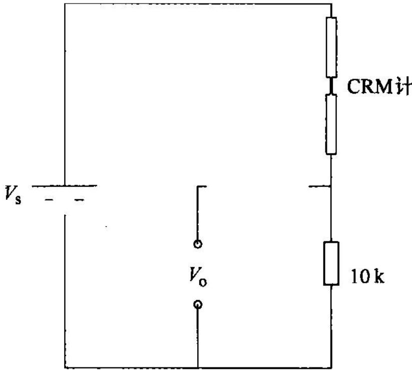


图 21-8　使用 CRM 计的电路

如果被嵌入的一个无源 RFID 标签中（见图 21-9），CRM 计的状态可能会被附近的 RFID 读卡器质疑。RFID 标签由一个天线环和一个微型芯片组成，后者包含

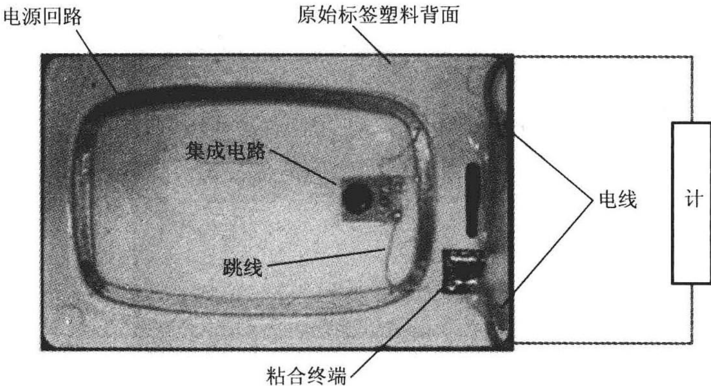


图 21-9　把计嵌入到被动 RFID 标签中

少量的信息（相关的 ID，例如结构的本地应变计），当标签被读卡器激励时会从标签发送给读卡器。注意，这种方法中，标签从读卡器获得能量并且它自身工作不需要能量。应变计用作天线环和微芯片的开关。如果应变计被折断，读卡器将不能与微芯片通信，预示存在的裂缝和变形会引起超过被选门槛的全部维度变化。

CRM 计也已经用于传感器电路，如图 21-8 所示。在这个配置中，应变计折断之前输出等于 V_s。折断后，输出等于零，表明存在过度的裂缝或者变形。

21.6　数据获取和通信

嵌入在建筑里的应变计可以给建筑的健康和任何出现的结构损坏一个准确的视图。应变计的状态必须是连续地监测和转换成适合计算机操作的数字形式。这需要使用数据获取设备，这些设备能够很容易地与应变计相结合。此外，获取的数据必须发送给远处的控制中心进行及时的分析。我们寻求一个灵活的、有序的无线解决方案进行通信，因为监测的建筑或者它们所在地区可能没有有线网络连接，并且安装大规模的有线网络是昂贵的、易受破坏的。

21.6.1　无源 RFID 设计

设计报告裂缝（21.5 节）的 CRM 计可以由无源 RFID 支持。这种情况下应变计的状态是一个二进制的值：断开或者不断开。我们把应变计嵌入到一个无源 RFID 标签中，如图 21-9 所示。嵌入应变计到标签天线回路并包含一个必要的段来完成数据报告电路。当应变计被断开时，电路打开并且标签不会回应 RFID 读卡器的质疑。否则，标签将会在质疑时回应它的 ID。被监测建筑的裂缝可以通过映射非响应标签的位置被调查。

无源 RFID 标签的优势是标签的成本非常低（每片低于 1 美元），尤其是大量订购时。除此之外，它们不需要能量来工作，因为它们感应来自响应的 RFID 读卡器的能量来工作。局限性是与有源 RFID 相比无源 RFID 有一个相对短的读取距离。为了分期偿还读卡器的费用（50 美元以上），使用一个读卡器来控制一系列标签，所有的这些标签必须相互接近。而且，读卡器/标签的实现必须支持 MAC 协议来阻止读卡器和标签之间信号的干扰，例如时障终止自适应收集协议［MIT03］。

为了把 RFID 数据发送给控制中心，我们把读卡器的通信界面连接到一个无线设备。对于最初建筑里转发数据，即室内环境，可以使用诸如 802.11 无线网络或者蓝牙技术。两种通信方式有不错的几十到几百米的通信范围，整个网络连接不需要太多通信节点就可以形成无线连接。而且，它们的信号可以穿过诸如墙壁和家具等障碍，进一步避免了对室内通信丢失的顾虑。特殊情况下，读卡器自身将会形成网络分区，无线中继节点可用于连接分区，尽管这需要事先规划或者现场调查。注意，某些读卡器可能不配备无线接口。例如，使用在［CRM08］中的 Phidgets 读卡

器有一个USB端口进行通信。在那种部署中，多达4个Phidgets读卡器通过以太网端口被连接到Keyspan US-4A USB服务器，这反过来与一台Linksys无线路由器进行无线连接。因此，实施细节变得更复杂，但概念设计保持不变。

建筑内转发的数据将最终需要从建筑外发送给控制中心。在户外环境，一个无线网络的节点通常为延伸的范围安装定向或者全向天线，它提供的广域无线传输是成熟且相对容易测试的技术。这样的网络运行在一个Ad hoc的802. 11x模式。没有固定基础配套设施，它采用了Ad hoc路由协议，例如AODV，用无线路由的具体指标（如ETX）发现数据接入点之间的对等高质量路径。

21.6.2　节点的设计

MICA2节点也可以用来质疑应变计的状态。这些节点可以购买于各种商业传感器，例如加速计、温度计、磁力计和测斜仪。然而，在我们的方案中传感器就是CRM计。数据收取板，例如MDA320CA或者ADAM-6017，在这个方案中需要连接到MICA2的处理器和无线模块。在MDA320CA的方案中，多达8个应变计可以分别连接到它的ADC和数字I/O口。这样一来，应变计采用MDA320CA的传感器激励输出2.5V供电，并且该节点可以通过获取板的应变计状态电平。注意，因为这个方案中应变计被供电，所以CRM计和普通电阻计都可以使用。

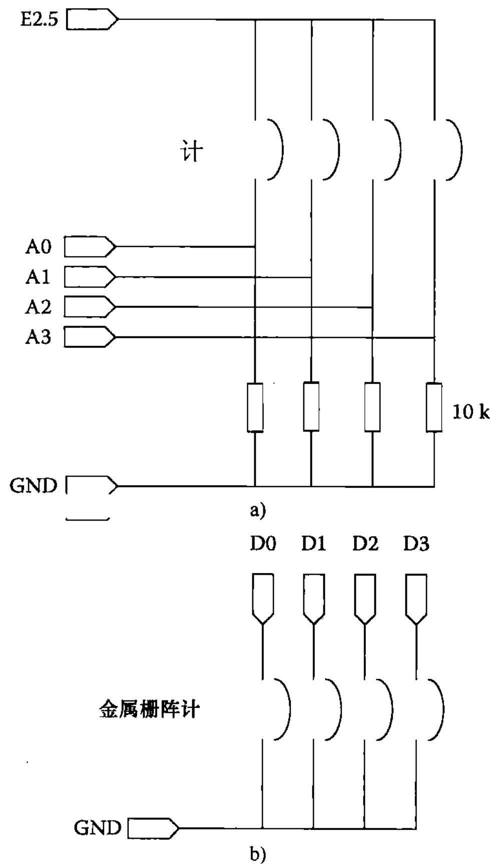


图21-10　数据获取板的布线计

图21-10a显示了四个CRM计写入了板的单端电压ADC输入。E2.5终端指出了板上2.5V的激励信道，并且这个终端连接到每个应变计的一端。应变计的另一端连接到ADC信道的输入以及10kΩ的下拉电阻。下拉电阻用来阻止ADC输入保持在浮动的状态，它的电压不确定并受到环境的干扰。这个接线在应变计没有断开时给一个2.5V的读数，当应变计断开时给一个0V

电平。图 21-10b 显示了应变计到数字信道的接线。接线连接应变计数字输入的一端，另一端到电气地。由于数据获取板内部的上拉电阻，应变计没有断开时接线给一个低的电平（GND）（因为应变计关闭了到 GND 的电路），断开时给一个高的电平（V_{CC}）。

一个两级的无线网络可以用来把数据传输给控制中心［CRM08］。低的一级用作应变计的多跳中继把室内数据报告给节点网关。室内通信以 MICA2 的 900MHz 以上的微波出现，这里节点形成了终止节点网关的 sink 树。路由协议从所有数据源（即带有应变计的节点）收集数据并把数据发送给 sink 节点。一般来说，sink 树可能有内部节点，其中 MICA2 节点完全用来数据发送（即它们没有附加计）。

sink 树收集的数据反过来由上级户外、广域传输的 802.11x 无线网络的网关发送到控制中心。为了在 RF sink 树网络和 802.11x 网络间架个桥，节点网关作为一个实现带有 MICA2 和 802.11x 路由的计算节点。802.11x 网络有一个作为网状网络的同一操作，这个网络描述于 21.6.1 节中的无源 RFID 设计。

21.7　控制软件

我们讨论了对传感器的部署，使用的是建筑内部监测支持数据获取和通信的设备。然后本地设备可以通过远程的 SHM 软件系统进行配置和控制，形成一个具有特定监测参数的综合部署。［CRM08］内基于组件的 SHM 软件设计成便携可配置的，能够容纳各种硬件确保它们协同工作的能力。特别的是，它支持无源 RFID 和节点的部署，描述于 21.6 节中。

21.7.1　安装和配置传感器

SHM 软件发布一个单独的软件包，能够安装运行 NET2.0 框架的常见 Windows，OS/X，Linux 和 UNIX 平台。一旦被安装，系统就可以根据特定的应用进行配置。这通过修改一个位于安装目录中的 XML 传感器配置目录 SensorConfig.xml 来实现。图 21-11 是一个配置目录的例子。配置的结构为一个 XML 基础树。根树是 sensorconfig，它的孩子是特定传感器类型的传感元件。每种传感器元件指的是传感器类型，并且各种同种型号的传感器实例可以由特定的 id 属性区分开。

传感器的详细操作由传感器元件进行配置，例如，就连接数据采集板的应变计而言，网络通信的协议和端口，应变计连接的板的数据信道，数据报告的轮询间隔，以及返回的报告读数说明。一般来说，XML 提出了每个传感器基于语义的描述，语义可以以特定的传感器总类进行定义。

21.7.2　实验配置

SHM 软件的核心是一套软件组件，进行设备/通信控制和传感器事件订阅，数

据采集，记录，选择和显示。进行实验的 SHM 部署然后配置为软件组件使用的图标。组件通过由完美定义的数据/控制界面控制的数据/控制流相互影响。

一个 SHM 的实验例子如图 21-12 所示。它有三层：应用层、通用传感层和实例传感层。应用层通过接收用户控制实验的指令和在可视界面显示建筑状态更新的视图与终端用户相互作用。通用传感层就传感器自身和传感信道（例如 21.6.2 节描述的 MDA320CA 数据信道）提供了一个抽象的传感器概念。抽象界面的实例方法包括打开和关闭传感器，识别传感器，质疑传感器的数据/错误状态，读取传感器数据，设置传感器的检测时间表等。实例传感层提供了对通用传感器和传感信道类的实际实施。它实现了传感器硬件需要的低层次通信协议执行由通用接口定义的高层次任务。例如，MODBUS 协议经常用作实际存在的包含 MICA2 数据采集板且带有广泛工业电子设备的通信标准。SHM 软件实现了 MODBUS 查询/响应消息的控制中心，消息通过 TCP/IP 与 MODBUS 传感设备进行通信。实例传感层整齐地封装了传感器硬件的详细驱动，以及用户隐藏的详细信息。

```
<?xml version="1.0" encoding="utf-8"?>
<sensorconfig>
  <sensor id="200" type="phidgetRFID">
    <description>Phidgets RFID Reader (17959)</description>
    <hardwareid>4627</hardwareid>
  </sensor>
  <sensor id="100" type="MDA320CA">
    <description>Crossbow MDA320 Data Acquisition Board
    </description>
    <port>TCP:192.168.1.6:10002</port>
    <channel id="0">
      <description>MDA320 Digital Channel #1</description>
    </channel>
    <channel id="1">
      <description>MDA320 Digital Channel #2</description>
    </channel>
  </sensor>
  <sensor id="300" type="ADAM6017">
    <description>ADAM6017 Data Acquisition Module
    </description>
    <port>TCP:192.168.1.160:502</port>
    <pollinterval>1000</pollinterval>
    <channel id="0">
      <description>ADAM6017 Channel #0</description>
      <range min="-5" max="5"/>
    </channel>
    <channel id="1">
      <description>ADAM6017 Channel #1</description>
      <range min="-5" max="5"/>
    </channel>
  </sensor>
</sensorconfig>
```

图 21-11　一个传感器配置文件的例子

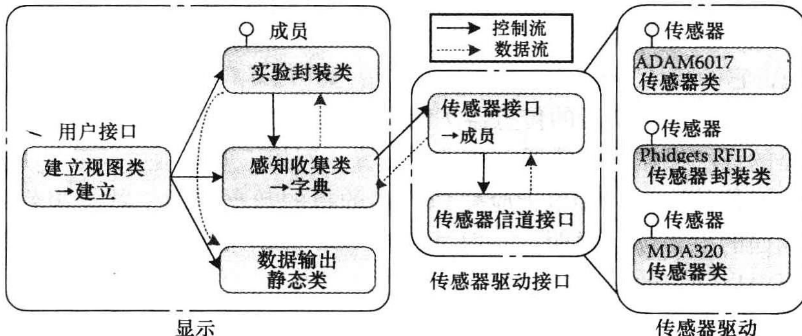


图 21-12　一个软件配置的例子

21.7.3 数据记录和显示

试验中，传感器采集的所有数据和事件被记录到一个 . edf 的记录文件里。记录文件包含实验的详细信息，包括传感器配置、实验开始/结束的数据/时间、所有收到的包括它们来源和时间的传感事件/数据报告，以及实验中发生的任何错误。

并不是所有详细记录的信息都是关于用户的切身利益。因此，图 21-12 中 DataExport组件的作用是根据用户的需求过滤记录数据并把数据安排成用户可读的形式。因为记录文件被写在了 XML 里，DataExport 可以实现一个定制的程序或者 Xquery 的 XSD 转换。这使终端用户不需要编程知识的情况下开发自己的数据输出功能，DataExport 对于诸如 Microsoft Excel 的大众化软件是完全兼容的。

21.8 CRM 计功能测试

基于 CRM 计监测系统的可行性评估了 CRM 计的特定配置。这个配置中，应变计包括两个各自长度为 $l_s/2$ 薄钢带和从一个地带跨越到另一地带的铜胶带(见图 21-13)。l_s 选择为 16in。钢带有 3/8in 宽，0. 005in 厚。使用的铜带是 GC 工具纯铜电路带（Tool Pure Copper Circuit Tape），有 1/8in 宽，0. 002in 厚。

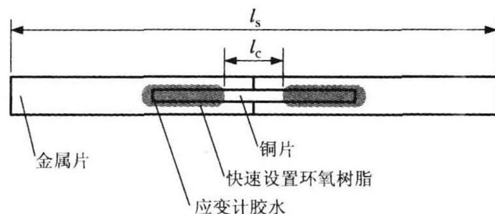


图 21-13 功能性测试中使用 CRM 计

图 21-13 显示的应变计用来检测受到不断地轴向载荷和循环横向载荷的圆柱的裂缝。这个柱子在日本名古屋的名古屋研究所（NIT）进行了测试（见图 21-14）。

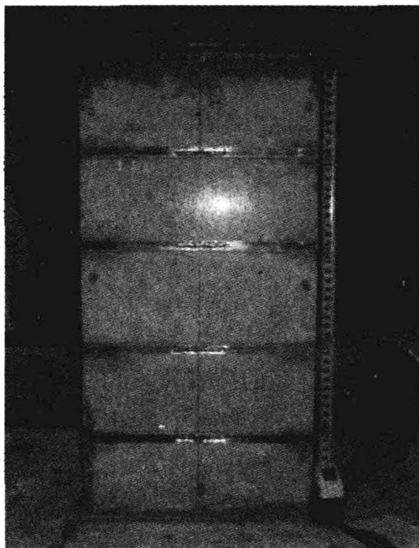


图 21-14 功能性测试中的实验装置：CRM 计部署采用银片

柱状实例中，四个垂直面的两个面各自安装了四个应变计：两个长度为 1/8in 的应变计和两个长度为 3/8in 的应变计。使用的两种应变计的位置沿着每个面交替出现的。这些应变计是面向图 21-14 所示的（垂

直于柱子的纵轴）。这种方向可以使应变计来检测纵向轴形成的夹角。这种类型的裂缝在过去的地震中被重复地观察到导致结构性故障。

原型计在 NIT 测试之前进行了校准，估计破裂变形和计长度之间的关系（见图21-15）。校准是通过在应变计上用两个不同长度的2in直径的混凝土圆柱实现的，应变计的各个端安装在不同的圆柱上。强度加大的钢栏浇铸在每个圆柱上。该钢栏用来在柱子上安装万能试验机器。使用这种机器，圆柱可以以恒定的速率被另一个所替换。分裂时圆柱间的总距离会被记录下来。裂缝可以通过传感器发射的电压信号检测到。信号通过使用数据采集系统被监测。裂缝通过传感器之间的两端电压突然变化感应到。

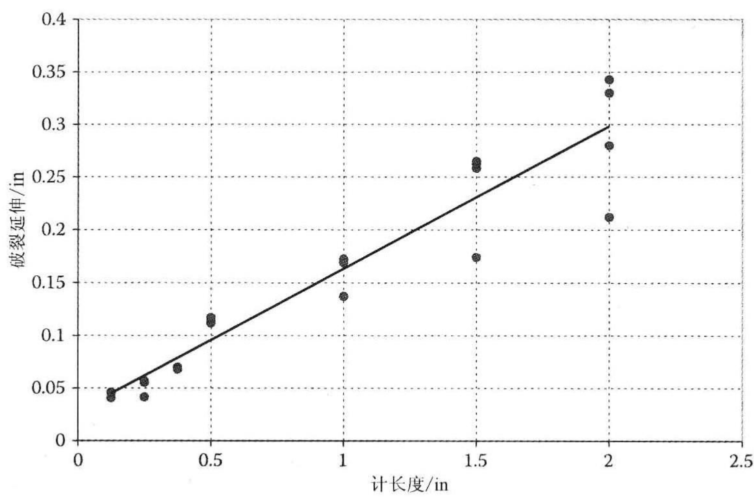


图21-15 破裂延伸与计长度 L_C 绘图

21.8.1 测试结果

测试的横跨直径的裂缝宽度可以使用裂缝宽度比较器进行测量。在测试期间，8个应变计中的6个监测到了足够大的裂缝可以在载荷完成前使应变计断裂。平均而言，长度为1/8in的应变计预计会在总变形为0.05in时断裂。长度为3/8in的应变计预计会在总变形为0.07in时断裂。表21-1列出了所观察的横跨应变计和产生断裂的总裂缝宽度。测量的总裂缝宽度匹配（平均）断裂的预期变形。

表 21-1　破损或者函数终端的延伸

计 ID	计长度/in	终端/破损的延伸		破碎?
		设计/in	实验/in	
1	1/8	0.050	0.050	Y
2	1/8	0.050	0.050	Y
3	3/8	0.070	0.075	Y
4	3/8	0.070	0.065	Y
5	1/8	0.050	0.050	Y
6	1/8	0.050	0.060	Y
7	3/8	0.070	0.070	N
8	3/8	0.070	0.080	N

21.9　大规模部署 CRM 计

全面的三层钢筋混凝土建筑结构 CRM 计的试验评估已经在 [CRM08] 中报道了。其结构作为关于平板型结构地震反应调查的一部分是由美国普渡大学 [Fick2008] 研究人员建立的。

结构的总高度是 30ft（1ft =0.3048m）。6 根钢筋混凝土柱，18 ×18in 的横截面积，支持了总共三个 7in 厚的平砖。每个板为 50ft，计划为 30ft。试验结构如图 21-16所示。其结构受到 $12ft^2$ 叠加充满水的桶的载荷。交替循环的载荷使用液压执行器应用在每个平板上。这些载荷引起了平板的断裂（见图 21-17）。我们实验的目的是评估 CRM 计检测这些裂缝的潜力。一套 4 个的 CRM 计部署在三个柱子各自的地板上，总共 12 个。每个 CRM 计配置来检测超过 0.03in 的累积裂缝。

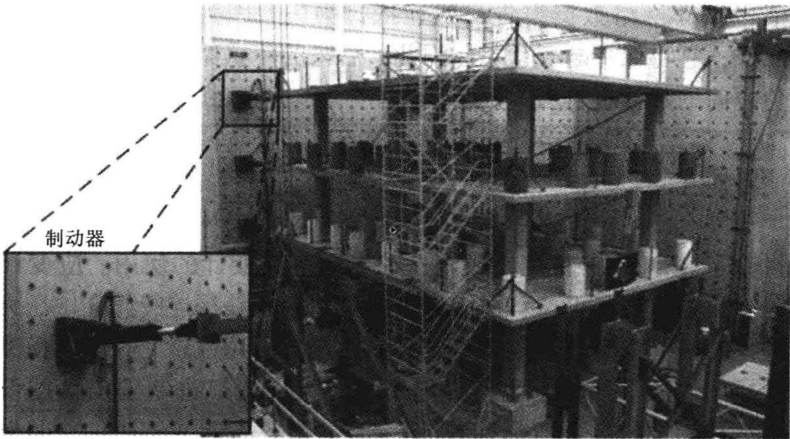


图 21-16　全面试验中的三层加强的混凝土建筑结构

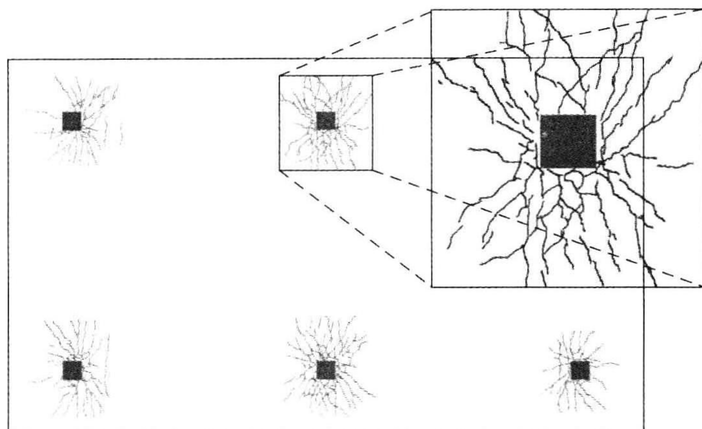


图21-17 平面图显示平板上的裂缝

在测试的不同阶段，由SHM网络报告的每个应变计的状态和裂缝的位置与宽带就会被记录下来。记录显示，在某种意义上说，该监测网络实现了百分之百的监测率，它准确地报告超过预选阈值而存在的裂缝。某种意义上说，如果拥有的累积裂缝宽度小于阈值裂缝宽度，没有计生成的裂缝断裂，那么没有假阳性。

根据是否采用无源RFID还是使用基于节点的数据采集，网络性能是各不相同的。就无源RFID而言，读卡器只有在它失去与RFID标签的联系时才进行报告，因为标签的数据电路被裂缝断开了。对标签流失事件的可靠性报道通过使用TCP/IP得以保证。因此，网络数据率是非常低的。就基于节点的ADAM-6017采集数据而言，应变计的状态每2s报告一次，形成了每采集模块1.208KB的数据率。这种情况下，SHM网络数据报告自然地抵御损失，因为断开的计是稳定的事件，一旦计被断开，它将保持断开状态。因此，后来的传感器数据报告包括之前的数据报告，以免之前的数据报告丢失。

21.10 结论

我们讨论了SHM中各种惠斯顿电桥和电阻基应变计的使用。应变计显示施加压力的电阻变化。惠斯顿电桥和电阻应变计可以用来获得连续媒介中压力的精确测量值。这种使用需要稳定、准确的数据采集系统的可用性。它仅限于横穿应变计的裂缝不能形成的情况。然后，我们提出了一种大应变二进制输出计的设计，CRM计用于检测重要构件的巨大变形和裂缝。CRM计是基于Wood和Neikirk [WN01] 以及Morita和Noguchi [MN06] 的想法，通过建立感应硬件的电路分裂来产生一个

大的输出，与损害有明确的联系且不需要进行分析。

为了支持远程结构监测的作用，我们提出了使用低成本 COTS 数据采集/通信设备的多跳无线 SHM 网络的设计和实施。我们也提出了基于组件的 SHM 控制软件，远程配置和控制各种感应/通信硬件以及管理传感器报告的数据。控制软件组件本身可以以即插即用的方式实现不同的应用。它们允许用户选择感兴趣的数据报告和格式化数据进行演示。

我们报告了一个 CRM 传感器的功能性测试，显示它可以人工检测超过配置阈值裂宽的可靠裂缝。除此之外，我们报告了在全面三层钢筋混凝土建筑结构上部署 12 个 CRM 计的结果。这些结果再次显示传感器可以可靠地检测模仿的现实环境中超过预设阈值的裂宽。CRM 传感器报告的简单性允许以较低的网络带宽和较高的恢复丢包能力进行有效的实时结构监测。

参考文献

[CRM08] Chin J. C., Rautenberg J. M., Ma C. Y. T., Pujol S., and Yau D. K. Y., A low-cost, low-data-rate rapid structural assessment network: Design, implementation, and experimentation. In *Proceedings IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, Atlanta, GA, September 2008.

[DAL93] Dally J. W., Riley W. F., and McConnell K. G., *Instrumentation for Engineering Measurements*, 2nd edn., Wiley, New York, 1993.

[Fick08] Fick D., Testing and structural evaluation of a large-scale three-story flat plate, Doctoral dissertation, Purdue University, West Lafayette, IN, April 2008.

[KPC07] Kim S., Pakzad S., Culler D. E., Demmel J., Fenves G., Glaser S., and Turon M., Health monitoring of civil infrastructures using wireless sensor networks. In *Proceedings ACM/IEEE Information Processing in Sensor Networks*, Cambridge, MA, 2007.

[MG07a] Measurements Group Inc., Strain gage selection: Criteria, procedures, recommendations, Tech. Note TN-505-4, 16p, 2007.

[MG07b] Measurements Group Inc., Errors due to transverse sensitivity in strain gages, Tech. Note TN-509-4, 9p, 2007.

[MG07c] Measurements Group Inc., Strain gage thermal output and gage factor variation with temperature, Tech. Note TN-504-1, 13p, 2007.

[MG07d] Measurements Group Inc., Errors due to Wheatstone Bridge nonlinearity, Tech. Note TN-507-1, 5p.

[MG07e] Measurements Group Inc., Shunt calibration of strain gage instrumentation, Tech. Note TN-514, 19p, 2007.

[MIT03] MIT Auto-ID Center, 3.56 MHz ISM band class 1 radio frequency identification tag interference specification: Candidate recommendation, verson 1.0.0. Technical report MIT-AUTOID-WH-002, 2003.

[MN06] Morita K. and Noguchi K., Crack detection sensor using RFID-tag and electrically conductive paint. Building Research Institute, Tsukuba, Japan, 2006.

[WN01] Wood S. L. and Neikirk D. P., Development of a passive sensor to detect cracks in welded steel construction. U.S.–Japan Joint Workshop and Third Grantees Meeting, Seattle, WA, 2001.

[XRC04] Xu N., Rangwala S., Chintalapudi K. K., Ganesan D., Broad A., Govindan R., and Estrin D., A wireless sensor network for structural monitoring. In *Proceedings ACM Sensys*, New York, 2004.

国际视野　科技前沿

国际信息工程先进技术译丛

《RFID与传感器网络：架构、协议、安全与集成》
《信号统计分析方法 —— 生物医学和电气工程应用指南》（原书第3版）
《数值方法在生物医学工程中的应用》
《生物医学工程学概论》（原书第2版）
《电生理学方法与仪器入门》
《医疗电子仪器的设计与开发》
《第三代移动网络中的多播通信：服务、机制、性能》
《电力线通信技术与实践》
《现代通信原理》（原书第2版）
《认知无线电网络》
《高速数字系统的信号完整性和辐射发射》
《UMTS中的LTE：基于OFDMA和SC-FDMA的无线接入》
《全面的功能验证：完整的工业流程》
《无线Mesh网络架构与协议》
《UMTS蜂窝系统的QoS与QoE管理》
《半导体制造与过程控制基础》
《WCDWA原理与开发设计》
《下一代移动系统：3G/B3G》
《IMS:IP多媒体概念和服务》（原书第2版）
《下一代无线系统与网络》
《深入浅出UMTS无线网络建模、规划与自动优化：理论与实践》
《HSDPA/HSUPA技术与系统设计——第三代移动通信系统宽带无线接入》
《无线传感器及元器件：网络、设计与应用》
《印制电路板——设计、制造、装配与测试》
《IPTV与网络视频：拓展广播电视的应用范围》
《多电压CMOS电路设计》
《微电子技术原理、设计与应用》
《蜂窝网络高级规划与优化2G/2.5G/3G/…向4G的演进》
《基于蜂窝系统的IMS——融合电信领域的VoIP演进》
《无线网络中的合作原理与应用》
《移动电视：DVB-H、DMB、3G系统和富媒体应用》
《环境网络：支持下一代无线业务的多域协同网络》
《基于射频工程的UMTS空中接口设计与网络运行》
《未来UMTS的体系结构与业务平台：全IP的3G CDMA网络》
《UMTS-HSDPA系统的TCP性能》
《宽带无线通信中的空时编码》
《数字图像处理》（原书第4版）
《基于4G系统的移动服务技术》
《吉规模集成电路互连工艺及设计》
《高性能微处理器电路设计》

CRC Press
Taylor & Francis Group





上架指导　工业技术／信息工程

ISBN 978-7-111-37399-5　定价：138.00元